AD_____

Award Number: DAMD17-99-1-9252

TITLE: Repair Machinery for Radiation-Induced DNA Damage

PRINCIPAL INVESTIGATOR: David Wilson, Ph.D.

CONTRACTING ORGANIZATION: Lawrence Livermore National Laboratory
Livermore, California 94550

REPORT DATE: July 2000

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 074-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	July 2000	Annual (1 Jul 99 - 30 Jun 00)

4. TITLE AND SUBTITLE
Repair Machinery for Radiation-Induced DNA Damage

5. FUNDING NUMBERS
DAMD17-99-1-9252

6. AUTHOR(S)
David Wilson, Ph.D.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Lawrence Livermore National Laboratory
Livermore, California 94550

E-MAIL:
wilson61@llnl.gov

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES
Report contains color photos

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 Words)*
Understanding the repair mechanisms for ionizing radiation (IR)-induced DNA damage and having prior knowledge of a patient's radiation-specific repair capacity will help determine which patients will be most responsive to radiation therapy and design more effective treatment regimes. The objective of this work has been to define the contributions of the mammalian protein Ape1, and other candidate nucleases, to the repair of IR-induced genetic damage. We are currently constructing cell lines that lack Ape1 protein and will determine the sensitivity of these mutant cells to various DNA-damaging agents, e.g. IR. In the course of these studies, we generated a mammalian cell line that overexpresses Ape1 roughly 5 to 8-fold. This line, relative to the non-overexpressing controls, will be used to examine whether Ape1 operates as a rate-limiting factor in the repair of certain exogenously-induced DNA damages (e.g. abasic sites and 3'-phosphate and phosphoglycolate blocking damages). Such studies will address the potential value of modulating Ape1 repair capacity in terms of protecting or sensitizing target cells from the effects of IR anti-cancer treatments. Additionally, we have completed the biochemical characterization of two protein factors, human Sfn and Hem45/Isg20. Results from these studies indicate that neither protein has a substrate specificity that is consistent with a major role in repairing IR-induced DNA damage. It thus appears that Ape1 is the predominant (and perhaps only) repair factor for 3'-damages, although alternative less prominent corrective mechanisms must exist. Future studies will focus more on obtaining accurate quantitative determinations of the overall contribution of Ape1 to 3'-damage repair and IR protection and less on the search for alternative 3'-repair mechanisms.

14. SUBJECT TERMS	15. NUMBER OF PAGES
Breast Cancer	52
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Unlimited

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. Z39-18
298-102

FOREWORD

Opinions, interpretations, conclusions and recommendations are those of the author and are not necessarily endorsed by the U.S. Army.

___ Where copyrighted material is quoted, permission has been obtained to use such material.

___ Where material from documents designated for limited distribution is quoted, permission has been obtained to use the material.

___ Citations of commercial organizations and trade names in this report do not constitute an official Department of Army endorsement or approval of the products or services of these organizations.

N/A In conducting research using animals, the investigator(s) adhered to the "Guide for the Care and Use of Laboratory Animals," prepared by the Committee on Care and use of Laboratory Animals of the Institute of Laboratory Resources, national Research Council (NIH Publication No. 86-23, Revised 1985).

X For the protection of human subjects, the investigator(s) adhered to policies of applicable Federal Law 45 CFR 46.

N/A In conducting research utilizing recombinant DNA technology, the investigator(s) adhered to current guidelines promulgated by the National Institutes of Health.

N/A In the conduct of research utilizing recombinant DNA, the investigator(s) adhered to the NIH Guidelines for Research Involving Recombinant DNA Molecules.

N/A In the conduct of research involving hazardous organisms, the investigator(s) adhered to the CDC-NIH Guide for Biosafety in Microbiological and Biomedical Laboratories.

28 July 2000

PI - Signature Date

Table of Contents

Proposal Title: Repair Machinery for Radiation-Induced DNA Damage

Principal Investigator: David M. Wilson III, Ph.D.

Biology & Biotechnology Research Program, L-441
Lawrence Livermore National Laboratory
Regents of the University of California
7000 East Avenue
Livermore, CA 94551-0808

Office: (925) 423 0695
Fax: (925) 422 2282

INTRODUCTION

Current methods to treat breast cancer and preserve normal tissue involve total mastectomy followed by radiotherapy. The use of postoperative radiotherapy however has produced mixed results. While some randomized studies indicate that there is no major benefit from postmastectomy radiation treatment, there exists a subgroup of patients that benefit from this type of strategy. Thus, it is imperative that we design effective prognostic tools to identify patients that will profit, as well as those that may be harmed, by adjuvant radiotherapy. Radiation is cytotoxic to cells in part because it introduces lethal genetic damage. Thus, understanding the repair mechanisms for radiation-induced DNA alterations and having prior knowledge of a patient's radiation-specific repair capacity will help design more effective treatment regimes and help determine which patients will be most responsive to radiation exposure. We intend to identify and characterize factors involved in correcting common radiation-induced damages. Such studies will provide the necessary tools (both specific DNA repair genes and biochemical assays) for predicting an individual's repair capacity and thus potential radiation sensitivity (genetic predisposition).

BODY

Objectives:

1. **To construct an *APE1* knockout cell line (months 1-24) and examine the role of Ape1 in DNA repair (months 25-36).**

Radiation induces an array of DNA damages, including abasic lesions and strand breaks harboring 3'-blocking termini such as phosphoglycolate and phosphate groups (1,2). The human Ape1 protein has been shown to incise at AP sites and remove a subset of 3'-damages, as well as to stimulate the DNA-binding activity of several oncoproteins (e.g. p53, Fos and Jun) *in vitro* (3,4,5,6). Yet despite the basic understanding of the biochemical properties of Ape1, the *in vivo* function of the protein remains largely unclear, particularly as it relates to radiation protection. *APE1* knockout cell lines represent an essential step towards defining the biological contribution of this mammalian protein.

We predict that Ape1-deficient cells will demonstrate a quantitatively significant defect in the repair of certain DNA damages, but of which damages needs to be determined. We have selected Chinese Hamster Ovary (CHO) as the target cell line for knockout generation, since they present the many advantages described in the original proposal. Towards generating Ape1-null cells, we first determined the APE1 gene copy number in the CHO AA8 cell line. Fluorescence *in situ* hybridization with an APE1 bacterial artificial chromosome (BAC) genomic fragment, which we isolated from a CHO BAC library, revealed that APE1 exists as a single copy gene in AA8 (see Appendix, Figure 1); thus only a single targeting event is needed to generate a knockout line. As noted in the original proposal, we had previously obtained an APE1 CHO genomic clone spanning roughly 13.5 kb (graciously provided by Pablo Arenas, UTEP). This fragment has been partially sequenced and used to construct the targeting vectors, which are designed to delete a portion of the Ape1 C-terminal domain known to be essential for nuclease activity. In our first screen of 360 colonies for an APE1 knockout, we were unsuccessful in identifying a targeted deletion event. This result prompted us to consider the possibility that APE1 null cell lines are inviable. Thus, to prevent future complications of this sort, we have stably integrated a Tetracycline (Tc)-regulated human APE1 cDNA into the AA8 genome. Of roughly 32 selected stably-integrated clones, twelve were found to overexpress the human Ape1 protein at significant levels. Of these, one (termed AA8 human Ape1-3) was selected based on the criteria that it expresses Ape1 at the highest level in the absence of Tc (which is used to shut down human APE1 cDNA expression), and is most sufficiently turned off in the presence of Tc (see Appendix, Figure 2). When compared to AA8 cells alone (i.e. cells not harboring the Tc-regulated human APE1 cDNA), the AP endonuclease activity of the AA8 human Ape1-3 cell line is 5 to 8-fold higher in the absence of Tc; notably, this higher expression is more similar to what is observed in human HeLa cell extracts (see Appendix, Table 1). Importantly, the AA8 human Ape1-3 cell line exhibits stable human Ape1 protein expression for at least 6 weeks. Using this line, we performed another targeting screen with the original knockout vector, and again were unable to detect any targeted events in >700 independent clones. We have since designed a new targeting construct, assuming that the initial construct was problematic, and are currently performing analysis of newly transfected AA8 human Ape1-3 cells. The objective, once we have obtained an APE1-deleted AA8 human Ape1-3 cell line, is to add Tc to the medium, shutting off human Ape1 protein expression, and to monitor cell growth in the absence of significant abasic endonuclease activity. If such a function is essential for viability, cells will be unable to survive in the presence of Tc. As described in the original proposal, if cells do survive, we will examine for sensitivity of these abasic endonuclease-deficient cells to several DNA-damaging agents.

As a fortuitous product of this research, the AA8 human Ape1-3 cell line presents a unique tool to examine the role of Ape1 in the repair of various DNA damages as a rate-limiting factor. That is to say, since AA8 maintains roughly 5 to 8-fold less abasic endonuclease repair activity (see Appendix, Table 1), if this activity is rate-limiting in the repair of specific forms of DNA damage, then the AA8 human Ape1-3 cell line will exhibit improved resistance. Alternatively, the higher level of Ape1 protein expression may sensitize cells to certain DNA-damaging agents, by generating incised single-strand break intermediates that cannot be processed efficiently by subsequent pathway repair factors. Thus, results from these studies will test the feasibility of regulating Ape1 protein levels as a means of protecting or sensitizing target cells from the effects of anti-cancer agents such as ionizing radiation, bleomycin and the many alkylating compounds. These experiments present a new line of investigation not directly proposed in the original grant.

2. To characterize hYjeR and hXPMC2/Hem45 as 3' to 5' exonucleases and 3'-repair enzymes (months 1-36).

Ape1 has been shown to possess a poor 3'-repair activity on certain DNA substrates (3), thus alternative repair mechanisms presumably exist for such damages. We will search for additional 3'-damage processing activities, testing both the *APE1* knockout cells for this function as well as several recently identified human candidate 3' to 5' exonucleases.

Previous computer-based studies identified three proteins, XPMC2, ISG20/Hem45 and YjeR, which contain the structurally-conserved exonuclease motifs of the 3′-5′ exonuclease superfamily (7,8). Using traditional protein expression and purification methods, we have purified the human YjeR and ISG20/Hem45 proteins to near homogeneity. Substrate analysis has revealed that the human YjeR protein exhibits nuclease activity on primarily nucleic acid substrates of <5 nucleotides in length (see Appendix, Nguyen et al., in press). Thus, we have named the human protein Sfn, for small fragment nuclease. Such a substrate specificity indicates that Sfn is unlikely involved in repairing genetic damage on large chromosomal DNA, yet presumably plays a role in recycling short DNA fragments, such as those generated during repair excision events. We will not pursue further characterization of Sfn, since (given its substrate preference) this protein is unlikely to contribute significantly to ionizing radiation resistance.

Substrate analysis with ISG20/Hem45 found that this human protein degrades both RNA and DNA substrates, with DNA at roughly 50-fold less efficiency (see Appendix, Nguyen et al., in preparation). Furthermore, substrate specificity was selective for single-stranded nucleic acid fragments. These results suggest that this protein, like Sfn, is unlikely involved in resistance to ionizing radiation-induced DNA damage. Thus, we are not planning to pursue further characterization of ISG20/Hem45. As Objective 2 of the original proposal (i.e. to characterized Sfn and ISG20/Hem45 biochemically) will be completed in the next several months (much quicker than anticipated), we will instead focus our future efforts on constructing the APE1 knockout cells (Objective 1), characterizing the AA8 human Ape1-3 cell line, and identifying yet other alternative 3'-repair proteins or mechanisms if appropriate.

KEY RESEARCH ACCOMPLISHMENTS

- Completed characterization of the human YjeR equivalent, which we have termed Sfn for small fragment nuclease. This data will appear in an upcoming issue of *Journal of Biological Chemistry* (see Appendix for included pre-print).
- Completed characterization of the human Hem45 protein. Results from these studies are currently being organized for submission (see Appendix for included rough draft).
- Established an Ape1-overexpressing mammalian cell line (AA8 human Ape1-3) that will permit analysis of the role of this protein in DNA damaging agent resistance, particularly as a rate-limiting factor. This cell line permits regulation of Ape1 protein expression and is currently being used in the knockout targeting experiments in the event that a null cell is lethal.
- Have completed the biochemical characterization of the abasic endonuclease activities of the overexpressing cell line (AA8 human Ape1-3) relative to cell line control extracts.
- Constructed two targeting constructs for knocking out the APE1 gene in CHO cells. The first construct has been abandoned after screening over 1000 colonies for a targeting event without

success. Current screening is ongoing after transfection of the second targeting construct into the AA8 human Ape1-3 cell line.

REPORTABLE OUTCOMES

- Nguyen, L.H., J.P. Erzberger, J. Root and D.M. Wilson III. (2000) The human homolog of *E. coli* Orn degrades small single stranded RNA and DNA oligomers. *J. Biol. Chem.* In Press.

- Nguyen, L.H. and D.M. Wilson III. (2000) Human Hem45/Isg20 degrades single-stranded RNA and DNA *in vitro*. In preparation.

- Development of the AA8 human Ape1-3 cell line.

CONCLUSIONS

Understanding the repair mechanisms for ionizing radiation (IR)-induced DNA damage and having prior knowledge of a patient's radiation-specific repair capacity will help determine which patients will be most responsive to radiation therapy and design more effective treatment regimes. The objective of this work has been to define the contributions of the mammalian protein Ape1, and other candidate nucleases, to the repair of IR-induced genetic damage. We are currently constructing cell lines that lack Ape1 protein and will determine the sensitivity of these mutant cells to various DNA-damaging agents, e.g. IR. In the course of these studies, we generated a mammalian cell line that overexpresses Ape1 roughly 5 to 8-fold. This line, relative to the non-overexpressing controls, will be used to examine whether Ape1 operates as a rate-limiting factor in the repair of certain exogenously-induced DNA damages (e.g. abasic sites and 3'-phosphate and phosphoglycolate blocking damages). Such studies will address the potential value of modulating Ape1 repair capacity in terms of protecting or sensitizing target cells from the effects of IR anti-cancer treatments. Additionally, we have completed the biochemical characterization of two protein factors, human Sfn and Hem45/Isg20. Results from these studies indicate that neither protein has a substrate specificity that is consistent with a major role in repairing IR-induced DNA damage. It thus appears that Ape1 is the predominant (and perhaps only) repair factor for 3'-damages, although alternative less prominent corrective mechanisms must exist. Future studies will focus more on obtaining accurate quantitative determinations of the overall contribution of Ape1 to 3'-damage repair and IR protection and less on the search for alternative 3'-repair mechanisms.

REFERENCES

1. von Sonntag C (1987) The Chemical Basis of Radiation Biology. Taylor and Francis (London)
2. Ward JF (1988) DNA damage produced by ionizing radiation in mammalian cells: identities, mechanisms of formation, and reparability. *Prog Nucleic Acid Res Mol Biol* 35:95-125
3. Wilson DM 3rd, Takeshita M, Grollman AP, & Demple B (1995) Incision activity of human apurinic endonuclease (Ape) at abasic site analogs in DNA. *J Biol Chem* 270 (27): 16002-16007

4. Suh D, Wilson DM 3rd, & Povirk L (1997) 3'-Phosphodiesterase activity of human apurinic/apyrimidinic endonuclease at DNA double-strand break ends. *Nucleic Acids Res.* 25:2495-2500.
5. Xanthoudakis S, Miao G, Wang F, Pan YC, & Curran T (1992) *EMBO J.* 11:3323-3335.
6. Jayaraman, L., Murthy, K. G., Zhu, C., Curran, T., Xanthoudakis, S., & Prives, C. (1997) *Genes Dev.* 11:558-570.
7. Moser MJ, Holley WR, Chatterjee A, Mian IS (1997) The proofreading domain of *Escherichia coli* DNA polymerase I and other DNA and/or RNA exonuclease domains.*Nucleic Acids Res* 25:5110-5118
8. Koonin EV (1997) A conserved ancient domain joins the growing superfamily of 3'-5' exonucleases. *Curr Biol* 7:R604-R606
9. Izumi T, Hazra TK, Bodogh I, Tomkinson AE, Park MS, Ikeda S, & Mitra S (2000) Requirement for human AP endonuclease 1 for repair of 3'-blocking damage at DNA single-strand breaks induced by reactive oxygen species. Carcinogenesis 21:1329-1334.

APPENDICES

Figure 1. Fluorescence *in situ* hybridization analysis of CHO APE1. Hybridization of genomic DNA from an APE1-containing BAC clone with an AA8 metaphase chromosome spread indicates a single hybridizing region (seen in green; indicated by arrow). These results reveal that APE1 is present as a single copy in CHO AA8 DNA.

Figure 2. Western blot and incision activity analysis of protein extracts from AA8 hApe1-3 cells. AA8 hApe1-3 cells were either not exposed (-) or exposed (+) to 0.3 ug Tc/ml for 0 to 96 hr. The cells were harvested at the time indicated and lysed by sonication to prepare protein extracts. Protein concentrations were determined by Bradford assay. A) 10 μg of protein extract was electrophoresed on 12% SDS-polyacrylamide gel and then transferred to nitrocellulose and immunodetected using an anti-human APE1 antibody and chemilluminescence. "Ape1" represents 0.1 μg of purified human protein as the positive control. The immunoblot demonstrates that exposure to Tc shuts off human Ape1 expression (indicated by arrow) to nondetectable levels by 72 and 96 hr. B) 0.1 μg protein extract was used in incision assays as described by Erzberger & Wilson (1999 *J. Mol. Biol.* 290:447-457) with 1 pmol of an 18 base pair duplex DNA substrate harboring an abasic site in the center. Specific activity is reported as pmol of substrate converted from the full-length non-incised 18mer fragment to the incised 9mer product per minute per mg cell extract. Tc exposed cells exhibit a reduced level of abasic incision activity over 96 hr.

Nguyen, L.H., J.P. Erzberger, J. Root and D.M. Wilson III. (2000) The human homolog of *E. coli* Orn degrades small single stranded RNA and DNA oligomers. *J. Biol. Chem.* In Press.

Nguyen, L.H. and D.M. Wilson III. (2000) Human Hem45/Isg20 degrades single-stranded RNA and DNA *in vitro*. In preparation.

FINAL REPORTS

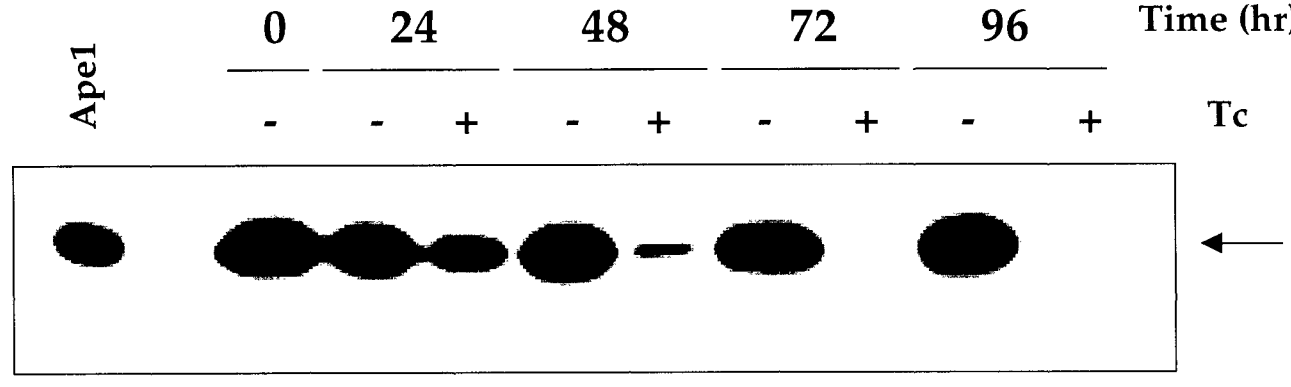
Nguyen, L.H., J.P. Erzberger, J. Root and D.M. Wilson III. (2000) The human homolog of *E. coli* Orn degrades small single stranded RNA and DNA oligomers. *J. Biol. Chem.* In Press.

Nguyen, L.H. and D.M. Wilson III. (2000) Human Hem45/Isg20 degrades single-stranded RNA and DNA *in vitro*. In preparation.

Role on Project	Name	Degree	Scientific Discipline	Institutional Affiliation
PI	David M. Wilson III	Ph.D.	Molecular Biology & Protein Biochemisty	Lawrence Livermore National Laboratory, University of California (LLNL-UC)
Post-doctoral Scientist	Laura Schild	Ph.D.	Molecular & Cellular Biology	LLNL-UC
Consultant	Larry Thompson	Ph.D.	Mammalian Genetics	LLNL-UC

Figure 1

A)



B)

	0 hr	24 hr	48 hr	72 hr	96 hr
- Tc	1812.2 ±11.52	1839.8 ±14.69	1835.9 ±17.49	1838.6 ±15.92	1830.7 ±23.67
+ Tc		1805.7 ±26.3	1016.3 ±338.3	383.5 ±12.3	266.0 ±40.0

Incision Activity (pmol/min/mg)

Figure 2

Specific Incision Activity of AA8 , AA8 hApe1-3 and HeLaS3 Cells

	AA8[1]		AA8 hApe1 –3[1]		fold increase over AA8	HeLa S3[2]		fold increase over AA8
Time (hr)	- Tc	+ Tc	- Tc	+ Tc	- Tc/+Tc	- Tc	+ Tc	- Tc/+Tc
0	218.7 ± 103.9		1812.2 ± 11.52		8.3	901.0 ± 352.7		4.1
24	272.3 ± 15.13	145.9 ± 15.6	1839.8 ± 14.69	1805.7 ± 26.3	6.8/12.3	914.8 ± 338.3	898.5 ± 326.7	3.4/6.2
48	251.2 ± 34.24	244.0 ± 52.3	1835.9 ± 17.49	1016.3 ± 338.3	7.3/4.2	1033.9 ± 533.7	746.9 ± 281.1	4.1/3.1
72	356.1 ± 170.38	153.8 ± 24.7	1838.6 ± 15.92	383.5 ± 12.3	5.2/2.4	719.7± 226.3	758.9 ± 274.6	2.0/4.9
96	209.5 ± 43.89	184.6 ± 39.5	1830.7 ± 23.67	266.0 ± 40.0	8.7/1.4	606.2 ± 276.3	1010.3 ± 321.9	2.9/5.4

Table 1. Specific Activity of AA8, AA8 hApe1-3 and HeLaS3 cell protein extracts. Cells were either not exposed (-) or exposed (+) to 0.3 μg Tc/ml for 0 to 96 hr. The cells were harvested at the time indicated and lysed by sonication to prepare protein extracts. Protein concentrations were determined by Bradford assay. 0.1 μg cell extract was used in the incision assay as described above in Figure 2 legend (Erzberger & Wilson 1999 *J. Mol. Biol.* 290:447-457). Specific activity is reported as pmol of substrate converted from the full-length non-incised 18mer fragment to the incised 9 mer product per minute per mg cell extract.

[1]Average of 4 experiments

[2]Average of 3 experiments

balt3/bc-bc/bc-bc/bc3300/bc7577-00a | stafford S=3 | 14/7/00 4:51 | Comments: | ARTNO: M002672200

THE JOURNAL OF BIOLOGICAL CHEMISTRY

Vol. 275, No. ??, Issue of ???? ??, pp. 1–xxx, 2000
Printed in U.S.A.

The Human Homolog of *Escherichia coli* ORN Degrades Small Single-stranded RNA and DNA Oligomers*

Received for publication, March 29, 2000, and in revised form, May 11, 2000
Published, JBC Papers in Press, June 12, 2000, DOI 10.1074/jbc.M002672200

Lam H. Nguyen, Jan P. Erzberger‡, Jeffrey Root, and David M. Wilson III§

From the Molecular and Structural Biology Division, Lawrence Livermore National Laboratory, L-441, Livermore, California 94551

We report here the identification of human homologues to the essential *Escherichia coli* Orn protein and the related yeast mitochondrial DNA-escape pathway regulatory factor Ynt20. The human proteins appear to arise from alternatively spliced transcripts, and are thus identical, except the human Ynt20 equivalent contains an NH_2-terminal extension that possesses a predicted mitochondrial protease cleavage signal. *In vitro* analysis revealed that the smaller human protein exhibits a 3′ to 5′ exonuclease activity for small (primarily ≤5 nucleotides in length) single-stranded RNA and DNA oligomers. We have named this human protein Sfn for small fragment nuclease to reflect its broad substrate range, and have termed the longer protein hSfnα. Sfn prefers Mn^{2+} as a metal cofactor and displays a temperature-resistant (to 50 °C) nuclease activity. Kinetic analysis indicates that Sfn exhibits a similar affinity for small RNAs and DNAs (K_m of ~1.5 μM), but degrades small RNAs ~4-fold more efficiently than DNA. Mutation of a conserved aspartate (Asp^{136}) to alanine abolishes both nuclease activities of Sfn. Northern blot analysis revealed that a 1-kilobase transcript corresponding to *SFN* and/or *SFNα* (these mRNAs differ by only two nucleotides) is expressed at varying levels in all fetal and adult human tissues examined. Expressed tag sequence clone analysis found that the two splice variants, *SFN* to *SFNα*, are present at a ratio of roughly 4 to 1, respectively. The results presented within suggest a role for human Sfn in cellular nucleotide recycling.

AQ: A

AQ: B

Nucleases are critical components of DNA and RNA metabolism, carrying out functions in DNA repair, replication and recombination, and RNA processing and degradation. Several candidate nucleases were recently classified as members of the 3′ to 5′ exonuclease superfamily (1). This superfamily includes RNases (such as RNase T and D), the proofreading domains of pol I family DNA polymerases, and DNases that exist as independent proteins (such as *Escherichia coli* Exo1) or as domains within larger polypeptides (such as a region within the Werner's syndrome helicase) (1–3). Homology within the superfamily is centered around three conserved exonuclease motifs (Exo I, II, and III) (2). The crystal structure of the Klenow fragment of *E. coli* DNA polymerase I and complementing site-directed mutagenesis studies indicate that the Exo I, II, and III motifs are clustered around the active site and contain four conserved negatively charged residues that are critical for coordinating the two metal ions involved in phosphodiester bond cleavage (reviewed in Ref. 4).

Based on conservation of the Exo motifs, the protein encoded by the *E. coli* open reading frame *YjeR* was placed into the 3′ to 5′ exonuclease superfamily (5). This protein was subsequently shown *in vitro* to be an exonuclease specific for RNA molecules shorter than about 5 nucleotides (nt)[1] in length (6). Based on this ribonuclease activity, *YjeR* was renamed *Orn* for oligoribonuclease (6). Data base mining revealed a human EST homolog to *YjeR*, indicating an evolutionary conserved role for the encoded protein (5).

Fn1

Orn is one of eight distinct 3′ to 5′ exoribonucleases present in *E. coli* (7). It is a processive enzyme that initiates attack at a free 3′ hydroxyl group on single stranded RNA, releasing 5′-mononucleotides in a sequential manner (6). Notably, unlike previously characterized *E. coli* ribonucleases, *ORN* is an essential gene. Experiments using a temperature-sensitive *ORN* construct and pulse-chase experiments with radiolabeled RNA revealed an accumulation of small RNA oligonucleotides at the non-permissive temperature (8). These observations suggested a role for Orn in the final steps of mRNA degradation, since mRNA-degrading enzymes such as RNase II and polynucleotide phosphorylase generate small oligonucleotide fragments that require further processing to mononucleotides (8).

An *ORN* homologue is also present in yeast *Saccharomyces cerevisiae* (5, 9). This *YNT20* gene encodes a larger protein that has been shown to localize to the mitochondria (9). *YNT20* plays a role in a *S. cerevisiae* mitochondrial DNA escape pathway, a process that involves the transfer of genetic material from the mitochondria to the nucleus (9).

It was suggested, based on sequence comparisons, that the *S. cerevisiae* Ynt20 protein is potentially capable of hydrolyzing both DNA and RNA (9). Although it has been reported that the *E. coli* Orn protein exhibits no detectable DNase activity on double stranded T7 chromosomal DNA (10), its ability to degrade short deoxyribonucleotides apparently has not been examined. We set out to examine the substrate specificity of the human Orn equivalent, with a particular interest in its activity for DNA molecules, and demonstrate here that this protein cleaves not only short RNAs but also short DNAs in a 3′ to 5′ direction. Thus, we suggest that the human enzyme be named Sfn for small fragment nuclease. We discuss our results as they may relate to the cellular role of this protein in nucleic acid metabolism.

* This work was supported by the U. S. Department of Energy by Lawrence Livermore National Laboratory under contract number W-7405-ENG-48 and supported by National Institutes of Health Grant CA79056 and U. S. Army Grant BC980514 (to D. M.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ Present address: Dept. of Molecular and Cellular Biology, University of California, Berkeley, CA 94720.

§ To whom correspondence should be addressed. Tel.: 925-423-0695; Fax: 925-422-2282; E-mail: wilson61@llnl.gov.

[1] The abbreviations used are: nt, nucleotide(s); EST, expressed sequence tag; PCR, polymerase chain reaction.

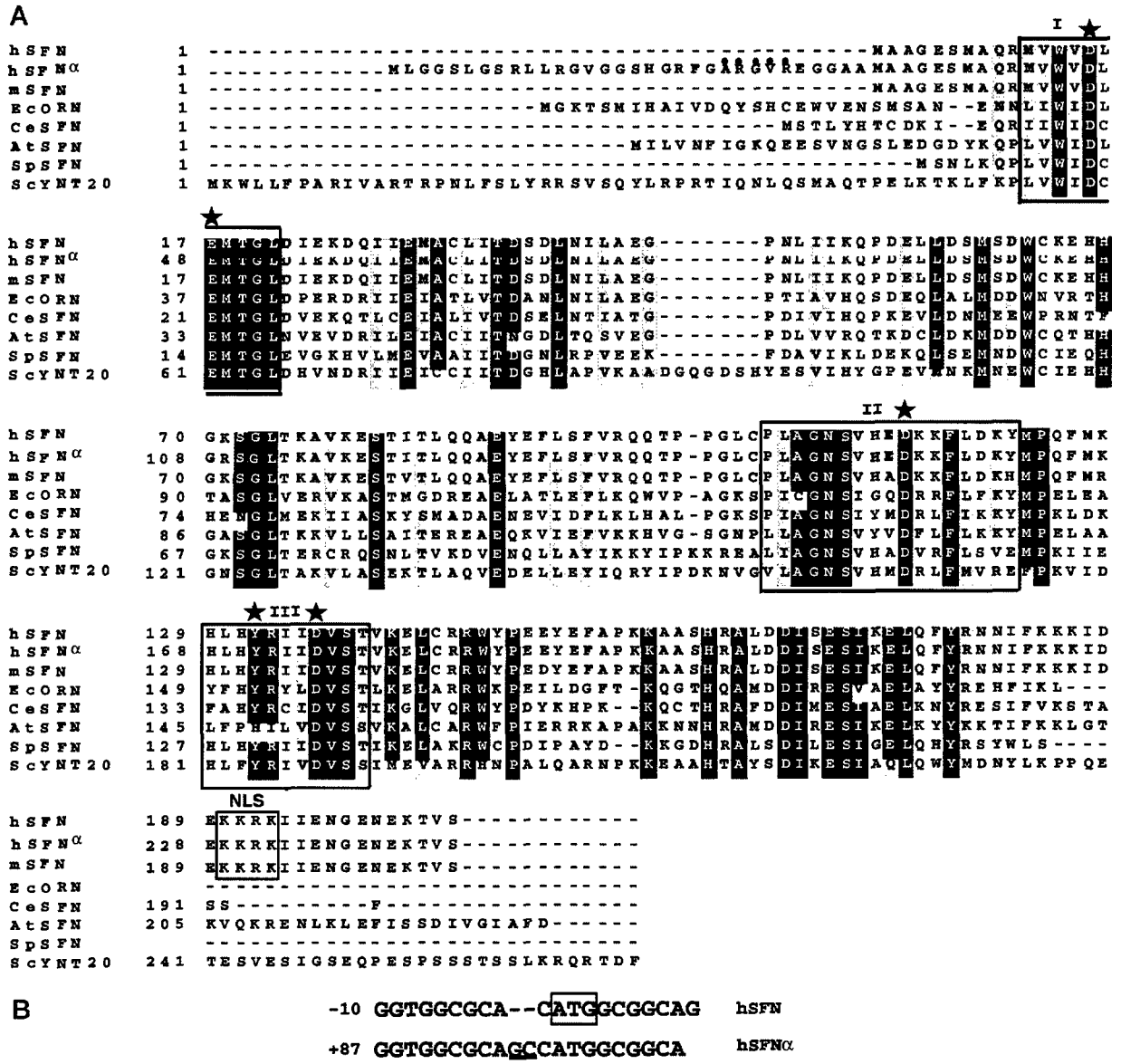


FIG. 1. **Comparison of amino acid and DNA sequences of the human *SFN*.** *A*, the amino acid sequence of the *SFN* subfamily. Amino acids in *black* are identical, *gray* are conserved, and *white* are non-conserved. This alignment was produced with the Boxshade program. *hSFN* is the AQ: F *Homo sapiens* Sfn protein sequence (accession number Q9Y3B8); *hSFNα*, *H. sapiens* Ynt20 (CAB53690); *mSFN*, *M. musculus* Sfn; *EcORN*, *E. coli* Orn (P39287); *CeSFN*, *Caenorhabditis elegans* Sfn (CAA96590); *AtSFN*, *Arabidopsis thaliana* Sfn (AAC77855); *SpSFN*, *Schizosaccharomyces pombe* Sfn (CAB37438); and *ScYNT20*, *S. cerevisiae* Ynt20 (P54964). The conserved exonuclease motifs of the 3′ to 5′ exonuclease superfamily (Exo I, II, and III) are shown as *boxes*. The *stars* indicate the catalytically important tyrosine and the four conserved negatively charged residues that are involved metal ion binding (1, 4). The nuclear localization signal (*NLS*), KKRK, is *boxed* (18). ● is used to mark the putative mitochondrial targeting signal, ARGVR, for *hSFNα* (17). *B*, sequences of the putative splice site in hSFN (accession number AF151872) and *hSFNα* (AL110239) mRNAs. The first nucleotide of the respective protein coding sequence is designated as +1. The first codon of *hSFN* is *boxed*. The differences in the two sequences are *underlined*.

EXPERIMENTAL PROCEDURES

Identification of SFN and Sfnα cDNAs—The EST data base was searched using the TBLASTN program to identify human cDNA clones that encode a protein homologous to the *E. coli* Orn protein (GenBank accession number P39287). IMAGE clone 663460 was chosen, sequenced (11), and found to contain an open reading frame encoding a 205-amino acid protein. This protein is identical to CGI-114 protein (accession number is AF151872). The *SFNα* cDNA was found in the GenBank data base (accession number AL110239).

Chromosomal Localization of the Human SFN Gene—The sequence at the 5′ end (accession number AA224296) or the 3′ end (accession number AA224194) of the *SFN* gene, when compared with the "Sequenced Tagged Site" data base at National Cancer Biology Institute, identified STS-H98169 as the genetic locus of *SFN*. This sequence-tagged site is positioned at chromosome 11 in a region between markers D11S1347 and D11S939 (110.3–117.9 centimorgan), which corresponds to the region 11q23.1–11q23.2. AQ: C

Buffers and Reagents—All reagents were purchased from Sigma unless otherwise indicated. Restriction enzymes were purchased from New England Biolabs. Labeled nucleotides were from Amersham Pharmacia Biotech. Spectrophotometric grade glycerol was obtained from Fisher. Olio(dT)$_{4-22}$ Ladder was purchased from Life Technologies, Inc. and consists of single stranded DNA oligos from 4 to 22 nt in length, increasing by 1-nucleotide increments. DNA oligos were obtained from Operon Biotechnologies (Alameda, CA). Synthetic RNA was obtained from Dharmacon Research (Boulder, CO). L buffer (Lysis buffer) con-

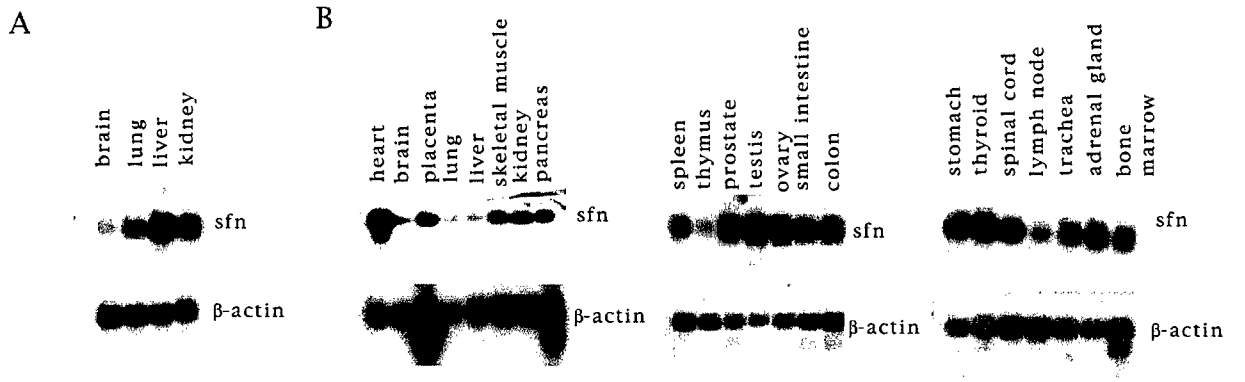


FIG. 2. **Expression pattern of *SFN* mRNA in fetal and adult human tissues.** Shown are Northern blot hybridization signals detected by autoradiography. Developmental stage, tissue type, and transcript type are indicated. β-Actin transcript was used for normalization. *Panel A* is an mRNA blot of fetal tissues. *Panel B* shows three mRNA blots of adult tissues.

tains 50 mM Tris-HCl, pH 7.9, 50 mM KCl, 20% glycerol, and 1 mM phenylmethylsulfonyl fluoride. W5 buffer is 50 mM Tris-HCl, pH 7.9, 500 mM NaCl, 20% glycerol, 5 mM imidazole, and 1 mM phenylmethylsulfonyl fluoride. W40 is 50 mM Tris-HCl, pH 7.9, 500 mM NaCl, 20% glycerol, 40 mM imidazole, and 1 mM phenylmethylsulfonyl fluoride. E250 buffer is 50 mM Tris-HCl, pH 7.9, 250 mM NaCl, 20% glycerol, 250 mM imidazole, and 1 mM phenylmethylsulfonyl fluoride. All pH values were determined at 21 °C.

Purification of Recombinant Sfn-His Protein—The *Sfn* coding region was PCR amplified using primers NCO5′HYJER (5′-GCATGCCTG-GCGGCAGGGGAGAGCAT) and HIND3′YJER (3′-GCAGTAAGCT-TACTCACGGTCTTCTCAT), and subcloned after digestion into the *Nco*I and *Hin*dIII sites of pET28d (Novagen, Madison, WI) to generate phyjeR-His. This construct allows for expression of six histidine residues on the carboxyl terminus under the control of a T7 RNA polymerase promoter. phyjeR-His plasmid was sequenced and no PCR errors were found.

The phyjeR-His plasmid was transformed into BL21(DE3)/pLysS *E. coli* strain (Novagen). An overnight culture of 100 ml was grown at 37 °C in LB (1% bactotryptone, 0.5% bactoyeast extract, and 1% NaCl) with 50 μg/ml kanamycin and 25 μg/ml chloramphenicol. The overnight culture was used to inoculate 1 liter of the same medium. This culture was grown at 37 °C with vigorous aeration until the $A_{550\ nm}$ was 0.8. Isopropyl-1-thio-β-galactopyranoside (final concentration of 1 mM) was then added to induce Sfn-His protein expression. At 4 h after induction, cells were harvested. The cell paste was resuspended in 30 ml of L buffer. The cell suspension was sonicated using a Misonix XL sonicator by three 1-min bursts with a microtip at the maximum setting. The cell lysate was centrifuged at 27,000 × *g* for 30 min at 4 °C. To the supernatant, polyethyleneimine was added slowly with constant stirring to a final concentration of 0.25% to remove nucleic acids (12). The suspension was centrifuged and the supernatant was loaded at 1.5 ml/min onto a 2-ml S2 cation exchange column equilibrated with L buffer using the BioLogic Workstation FPLC system (Bio-Rad). Fractionation over the cation exchange column removed the contaminating proteins found to bind nonspecifically to the subsequent Ni affinity column. The S2 flow-through was collected, and NaCl and imidazole were added to final concentrations of 500 and 5 mM, respectively. The S2 flow-through was then incubated with 2 ml of Ni-NTA resin (Qiagen, Santa Clarita, CA) for 1 h at 4 °C with gentle rocking. This suspension was centrifuged, and the resin was washed 2 times each with 20 ml of W5 buffer, followed by 4 washes each with 10 ml of W40 buffer. Sfn-His proteins were then eluted from the Ni-NTA resin 4 times, each with 2 ml of E250 buffer, dialyzed overnight against L buffer containing 0.1 mM dithiothreitol, and stored at −70 °C. Concentrations of protein solutions were determined by measuring the absorbance at 280 nm and using the theoretical molar extinction coefficient $E_{280\ nm}$ of 25,940 M^{-1} cm^{-1} calculated for Sfn-His (13).

Analytical Gel Filtration Chromatography—Sfn-His protein was separated on a 7.8 × 300-mm BioSil SEC 125–5 gel filtration column from Bio-Rad. The running buffer was 50 mM Tris-HCl, pH 7.9, 100 mM ammonium sulfate, 5% glycerol, 0.1 mM EDTA, and 0.1 mM dithiothreitol. The flow rate was 0.25 ml/min. Size markers used for calibration were thyroglobulin (670 kDa), γ-globulin (158 kDa), ovalbumin (44 kDa), myoglobin (17 kDa), and vitamin B-12 (1.35 kDa). Proteins were detected by ultraviolet absorbance at 280 nm.

Site-directed Mutagenesis of the SFN cDNA—Mutagenesis was performed using the overlapped PCR method as described (14). The site-specific mutant oligo used was D136A, 5′-TATAGAATAATTGCTGT-GAGCACTGTTAAG-3′, with the mutated codon underlined. We chose to mutate Asp^{136} because a structurally equivalent mutation in *E. coli* DNA pol I (D501A) caused a 13,000-fold decrease in 3′-5′ exonuclease activity (15). The final *SFN* mutant PCR product was digested with *Nco*I and *Hin*dIII and subcloned into the same restriction sites of pET28d. This construct was sequenced as described (11).

Nuclease Assays—Various DNA and RNA substrates were labeled at the 5′ end by T4 polynucleotide kinase with [γ-^{32}P]ATP (16). RNA5 is 5′-GAUCG-3′. DNA5 is 5′-GATCG-3′. DNA8 is 5′-CACGAGCC-3′. 100 fmol of nucleic acid substrate was incubated with 100–2000 fmol of Sfn-His protein at 37 °C in 10 μl of 50 mM K-HEPES, pH 7.4, 10% glycerol, 50 mM KCl, 10 mM $MnCl_2$ (unless otherwise noted), 0.01% Triton X-100, and 0.1 mM dithiothreitol. The reactions were stopped with 10 μl of 80% formamide dye solution, heated at 80 °C for 3 min, and then fractionated on a 20 or a 22.5% 8 M urea polyacrylamide gel. The gel was exposed to x-ray film. Visualization of the labeled substrate on the gels was also achieved using a Molecular Dynamics (Sunnyvale, CA) STORM 860 Phosphorimager and quantitative analysis was performed using Molecular Dynamics ImageQuant v1.11 software.

Kinetics of Sfn Exonuclease Activity on Single Stranded DNA and RNA Substrates—Nuclease assays were performed under standard buffer conditions as described above at 37 °C, with Sfn at a final concentration of 10 nM. The reaction was incubated for 5 min with RNA5 or 40 min with DNA5, and was within the linear range of enzymatic activity (*i.e.* ≤15% of the substrate was converted to product). The substrate concentration range was 0.1–3.2 μM. The apparent Michaelis-Menten constant (K_m) and maximal velocity (V_{max}) were obtained from a double-reciprocal Lineweaver-Burk plot. A linear plot of 1/[S] *versus* 1/*V* (where [S] is the substrate concentration in μM and *V* is the velocity in μM min^{-1} units) produces a slope of K_m/V_{max} and a *y* intercept of $1/V_{max}$. k_{cat} was calculated from the equation $(k_{cat})(E_t) = V_{max}$, where E_t is the total concentration of enzyme in the assay. Linear regression analysis was performed using CricketGraph software (Cricket Software, Philadelphia, PA).

Northern Blot Analysis—Northern blots were prehybridized for 3 h at 65 °C in ExpressHyb Solution (CLONTECH, Palo Alto, CA). The cDNA probe (Sfn PCR product described above) was labeled using the Megaprime DNA Labeling System (Amersham Pharmacia Biotech) and [α-^{32}P]dCTP (Amersham Pharmacia Biotech), and hybridized in ExpressHyb Solution at 65 °C for 2–3 h. Blots were washed once at room temperature for 30 min and twice at 50 °C for 30 min with 50 mM $NaPO_4$, pH 7.4, 0.5% SDS, and 1 mM EDTA. Images were exposed on Kodak Bio-Max MS film for 16 h and developed. Blots were normalized with β-actin transcripts.

RESULTS

The Sfn cDNA, Genomic Location, and mRNA Expression Pattern—A cDNA containing the *SFN* coding region was obtained from the EST data base (see "Experimental Procedures"). By blasting the sequence of the 5′ end (accession number AA224296) or the 3′ end (accession number AA224194) of the *SFN* cDNA against the NCBI UniGene data base, we found

that the human *SFN* gene mapped to chromosome 11 at position 11q23.1–11q23.2. Four additional genes also map within this vicinity: apolipoprotein A-I (*apoA-I*), human serotonin receptor 3 (*HTR3*), human nicotinamide *N*-methyltransferase (*NNMT*), and zinc finger protein ZNF259 (*PLZF*).

The *SFN* cDNA codes for a 205-amino acid protein of 23,754 daltons, with a theoretical pI of 5.6. *Sfn* belongs to the *yjeR/ORN* family, which is a distinct subgroup of the 3′ to 5′ exonuF1 clease superfamily (1, 5). Fig. 1 shows a comparison of the amino acid sequence of the Sfn-like proteins from bacteria, yeast, plant, worm, mouse, and human. The human Sfn protein is ~50% identical to its *E. coli* counterpart, the Orn protein (5). Sfn and *E. coli* Orn possess the three characteristic sequence motifs, termed Exo I, II, and III (2), of the 3′ to 5′ exonuclease superfamily (1) (Fig. 1*A*). Four conserved negatively charged residues within these three Exo motifs (shown in Fig. 1*A*) are involved in positioning of the two divalent cations required for catalysis and phosphodiester bond cleavage (reviewed in Ref. 4).

Upon further examination of the GenBank data base, an identical protein to hSfn, but with an extended NH_2-terminal domain, was identified (Fig. 1*A*). This protein, which we propose to call Sfnα, appears to be the human equivalent to the yeast Ynt20 protein and to have arisen from an alternative RNA splicing event where two additional nucleotides were introduced (Fig. 1*B*). Using the PSORT II computer search program, a putative consensus cleavage site motif for mitochondrial processing proteases (ARGVR) (17) was identified within the unique NH_2-terminal portion of Sfnα. Both Sfn and Sfnα contain a consensus nuclear localization signal (KKRK) (18) (Fig. 1*A*). It seems logical that Sfn would be targeted to the nucleus and that Sfnα would translocate predominantly to the mitochondria.

Northern blotting revealed that the *Sfn* cDNA probe detects a single transcript of ~1 kilobase in all fetal and adult tissues F2 examined (Fig. 2). Comparatively, relatively low mRNA levels are observed in adult lymph nodes, brain, lung, liver, spleen, and thymus, with highest levels observed in heart. Since the human *SFN* and *SFNα* cDNAs appear to differ by only two nucleotides, it was not possible to distinguish the two transcripts here. Of the 10 human ESTs found in the data base, 8 contained a sequence identical to the *SFN* cDNA and two maintained a sequence identical to SFNα.

Purification of Overproduced Sfn-His Fusion Protein—We have characterized here Sfn, and expect that Sfnα will exhibit similar substrate specificity as it maintains the same core nuclease domain (Fig. 1*A*). The *SFN* gene was subcloned into the pET28d plasmid to produce a COOH-terminal-tagged Sfn-His fusion protein, and the recombinant protein was purified as described under "Experimental Procedures." From 1 liter of induced bacterial culture (about 5 g of wet cell pellet), we obtained ~4 mg of Sfn-His fusion protein of >95% purity (Fig. F3 3*A*). As shown in Fig. 3*B*, the Sfn-His fusion protein migrates as a single symmetrical peak on an analytical gel filtration column that corresponds to a globular protein with a molecular mass of ~90 kDa. Since the calculated molecular weight of Sfn is 24,000, Sfn-His fusion protein appears to migrate as a tetramer. However, since the elution position on a gel filtration column is dependent on size and shape of the protein, it is possible that the Sfn-His fusion protein exists in solution as a rod-shaped protein of lower oligomeric state, since rod-shaped proteins are known to migrate slower than expected for their Fn2 molecular weight.[2] Unlike the human Sfn-His protein, the *E.*

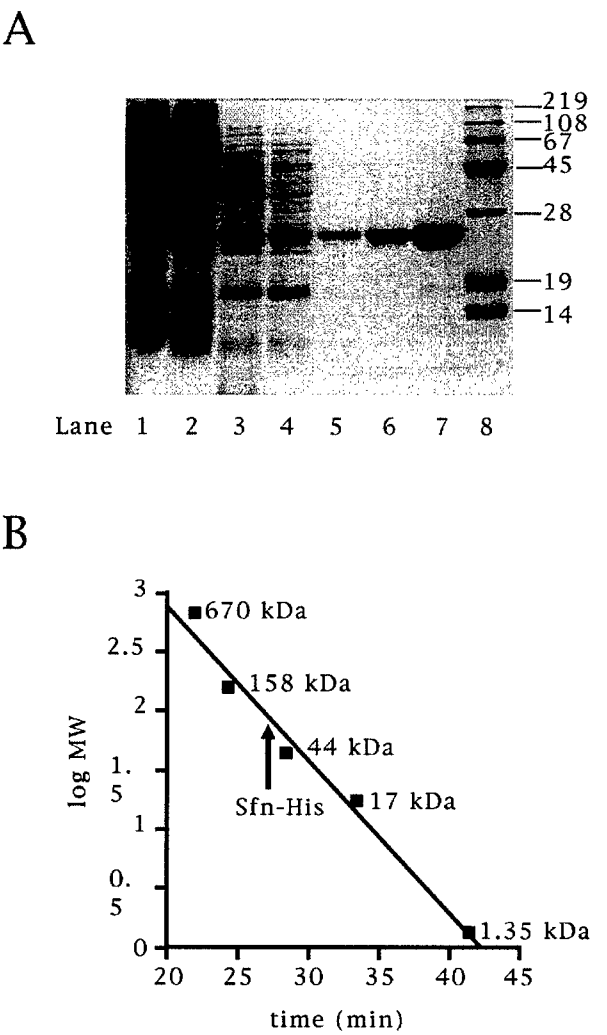


FIG. 3. **Purification of Sfn-His protein.** *A*, a 15% polyacrylamide SDS gel containing fractions of the various purification steps. *Lane 1* is 10 μl of whole cell extract (30 ml); *lane 2* is 10 μl of whole cell supernatant (30 ml); *lane 3* is 10 μl of the supernatant after 0.25% polyethyleneimine precipitation (30 ml); *lane 4* is 14 μl of the S2 cation exchange column flow-through (40 ml); *lanes 5–7* are 2.4, 7.2, and 21.6 μg of Ni-affinity purified Sfn-His (6 ml); and *lane 8* is protein markers. Protein marker sizes in kDa are indicated. The gel was stained with Coomassie Blue R-250. *B*, gel filtration column analysis of Ni-purified Sfn-His. The *x* axis is the retention time (in min) of marker proteins. The *y* axis is log of molecular weight (log MW) of proteins. The positions of the marker proteins are indicated on the calibration curve by their molecular weights. The retention time of Sfn-His is indicated by the *arrow* on the calibration curve.

coli Orn protein was reported to be a dimer based on its gel filtration chromatographic profile, although it should be noted that the experimental conditions differed slightly from those used here (6).

Exonuclease Activity of Sfn-His Protein—Short fragments of ≤5 ribonucleotides were shown to be the optimal substrate for *E. coli* Orn protein (6, 10). Using a 5-nucleotide RNA (RNA5), labeled at the 5′ end, as the substrate, we determined the divalent metal cation preference of the human Sfn. At metal concentrations of 0.01, 0.05, 0.1, 0.5, 1, 5, and 10 mM, only Mg^{2+} and Mn^{2+} had any stimulatory effect on Sfn-His exonuclease activity, with Mg^{2+} being ~10-fold less effective than Mn^{2+} at F4 10 mM (Fig. 4*A* and data not shown). Both human Sfn-His and *E. coli* Orn enzymatic activities have a strong dependence on

[2] Dr. R. Burgess, University of Wisconsin-Madison, personal communication.

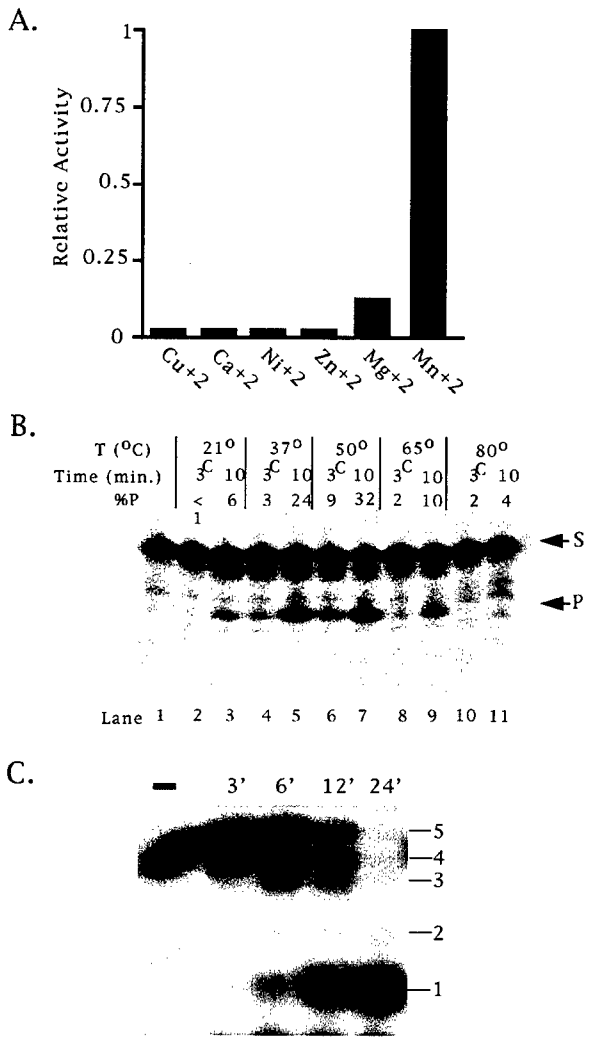


FIG. 4. **Metal cofactor requirement, temperature optimum, and time course of Sfn-His RNase activity.** *A*, metal preference of Sfn-His. 100 fmol of labeled RNA5 was incubated with 1 pmol of Sfn-His at 37 °C for 30 min in the presence of 10 mM of the different divalent metal cations as indicated. The *y* axis is relative exonuclease activity. Minimal detection for this assay was ≤1 fmol of product which is equivalent to ~2.5% of the activity detected with Mn^{2+} as shown for Cu^{2+}, Ca^{2+}, Ni^{2+}, and Zn^{2+}. *B*, temperature optimum of Sfn-His. 100 fmol of labeled RNA5 was incubated with 500 fmol of Sfn-His for 3 or 10 min at the temperature indicated. *Lane 1* is no protein control. *Lanes 2* and *3* represent reactions performed at 21 °C; *lanes 4* and *5*, 37 °C; *lanes 6* and *7*, 50 °C; *lanes 8* and *9*, 65 °C; and *lanes 10* and *11*, 80 °C. *%P* is the percentage of labeled substrate that has been converted to the final mononucleotide product. *S* is the labeled RNA5 substrate. *P* is the final product. Shown is a representative of three experiments. *C*, a time course of Sfn-His RNase activity with 5 mM $MnCl_2$ at 37 °C. *Lane 1* is the no protein control. *Lanes 2–5* are 0.2 μM RNA5 incubated with 0.02 μM Sfn-His for various times (in minutes) as indicated on *top* of the figure.

(and a preference for) Mn^{2+} (Fig. 4*A*; Ref. 10). The RNase activity of Sfn-His also has a temperature optimum around 50 °C (Fig. 4*B*), similar to that of *E. coli* Orn protein (10). Thus both proteins are quite thermostable. The RNA-specific degradation activity of Sfn at 37 °C is ~2-fold less than that at 50 °C (Fig. 4*B*). Since we observed degraded 4-mer products prior to detecting mononucleotide products (Fig. 4*C*), Sfn-His appears, as expected, to be a 3′ to 5′ exonuclease.

To determine whether short single stranded DNA can serve as substrate for Sfn-His, we used a single stranded oligo(dT)

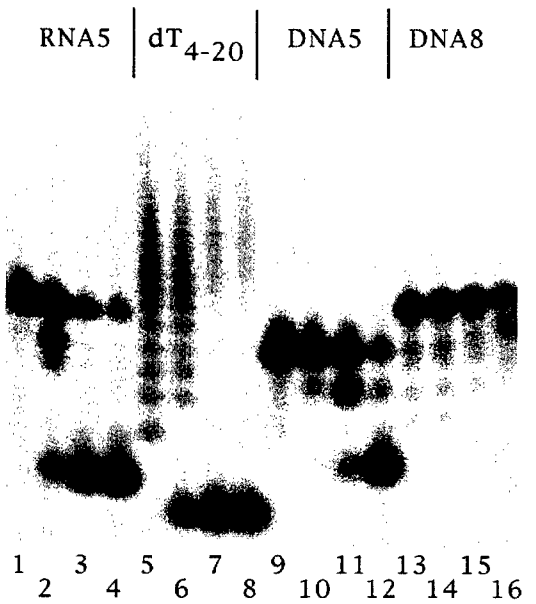


FIG. 5. **Sfn-His degrades small single stranded RNAs and DNAs.** All substrates were labeled at the 5′ end. 0.01 μM of each substrate was incubated with 0.2 μM Sfn-His protein for 1 h at 37 °C in the presence of 5 mM $MnCl_2$ for RNAs and 5 mM $MgCl_2$ for DNAs. Oligo(dT)$_{4-22}$ ladder consists of single stranded DNA oligos from 4 to 22 bases long, increasing by 1 base increments. DNA5 and DNA8 are single stranded DNA of 5 and 8 bases in length. *Lanes 1, 5, 9,* and *13* are the no protein controls. *Lanes 2, 6, 10,* and *14*, 5-min incubations with Sfn-His; *lanes 3, 7, 11,* and *15*, 15-min incubations; and *lanes 4, 8, 12,* and *16*, 45-min incubations.

ladder from 4 to 22 bases in length, and single stranded DNA oligos of 5 and 8 nt. As shown in Fig. 5, Sfn-His is capable of degrading short single stranded DNA, although with less efficiency than 5-mer RNAs. Furthermore, its DNase activity is inversely related to the length of the DNA oligo, as seen for the RNase activity of *E. coli* Orn (19). The data above is also consistent with the lack of DNase activity reported for Orn on double stranded T7 chromosomal (large) DNA (10).

The DNase and RNase Activities Are Intrinsic to Sfn-His—To confirm that the observed DNase and RNase activities are intrinsic to the Sfn-His protein, a catalytically inactive Sfn-His mutant was constructed. Amino acid residue Asp^{136} (Fig. 1, in Exo III motif) was selected because an aspartate to alanine mutation at the structurally equivalent residue (5) in *E. coli* DNA pol I causes a 13,000-fold decrease in 3′ to 5′ exonuclease activity (14). The D136A Sfn-His mutant was purified using the same procedure as described for wild type Sfn-His protein (Fig. 6*A*). When compared with wild type Sfn-His, D136A Sfn-His mutant displayed a ≥50-fold reduced exonuclease activity for both RNA5 and DNA5 substrates (Fig. 6*B*). These data are consistent with both nuclease activities being intrinsic to the Sfn-His protein, and not being the result of protein contaminants. The results also demonstrate the importance of this acidic residue in the enzymatic function of the 3′ to 5′ exonuclease family members (1).

We subsequently examined Sfn nuclease activity on DNA structures known to be products/substrates of DNA repair processes. As might be expected from the results above, in the presence of Mn^{2+} or Mg^{2+}, we did not detect any protein-dependent degradation of 42-base pair double-stranded DNAs of different configurations, including gapped, nicked, 4-nucleotide recessed or overhanged 3′ ends, or flap structures (data not shown; substrates described in Ref. 20).

Kinetic Parameters of the Sfn Nuclease Activities—To deter-

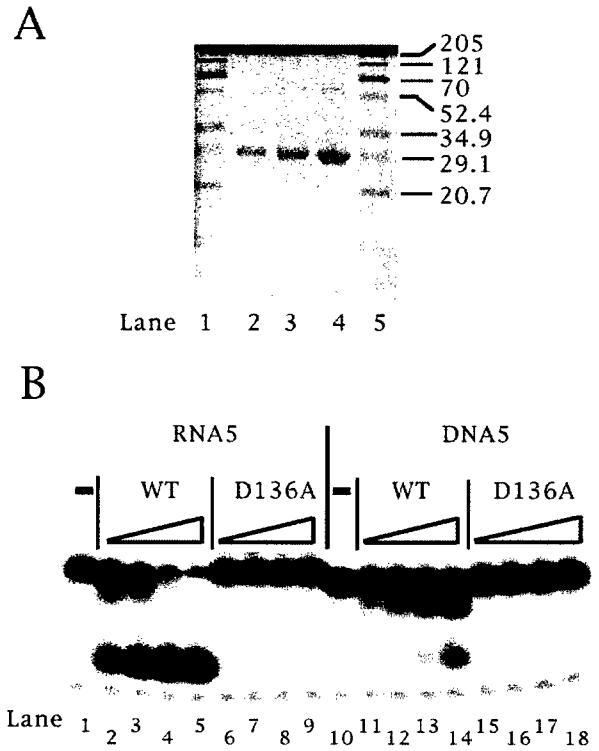


FIG. 6. **The effect of a D136A mutation on the DNase and RNase activities of Sfn-His.** *A*, purity of D136A Sfn-His mutant as determined on a 15% polyacrylamide SDS gel stained with Coomassie Blue R-250 dye. *Lanes 1* and *5* are protein markers; and *lanes 2–4* are 2.5, 5, and 10 μg of Ni-affinity purified D136A Sfn-His. Protein marker sizes in kilodaltons are indicated to the *right*. *B*, exonuclease activity of D136A Sfn-His mutant with RNA5 and DNA5 single stranded substrates. *Lanes 1–9* are with RNA5, and *lanes 10–18* are with DNA5 substrates. The substrate concentration is 200 nM. *Lanes 1* and *10* are the no protein control. *Lanes 2–5* and *11–14* are 12, 24, 48, and 96 nM of purified wild type Sfn-His (WT). *Lanes 6–10* and *15–18* are 12, 24, 48, and 96 nM of purified D136A Sfn-His mutant. The reactions were incubated at 37 °C for 20 min.

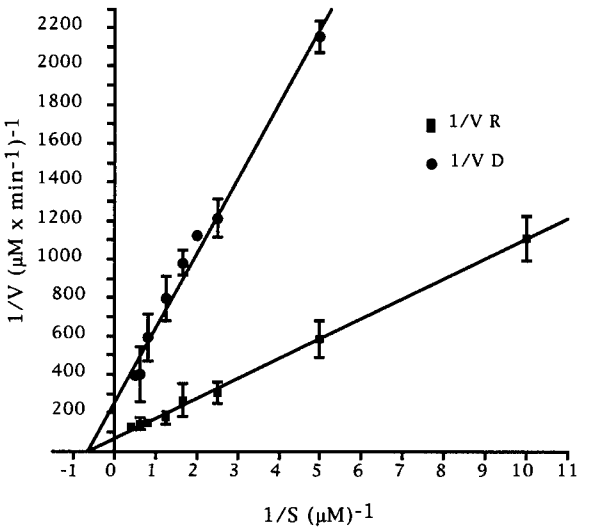


FIG. 7. **Double reciprocal Lineweaver-Burk plot of Sfn exonuclease activity on 5-nt single stranded RNA5 and DNA5.** The x axis is the reciprocal of substrate (1/S) in $(\mu M)^{-1}$ units. The y axis is the reciprocal of velocity (1/V) in $(\mu M \times min^{-1})^{-1}$ units. ■ represents RNA5; ●, DNA5 substrate. The data presented here is the average of five independent experiments for RNA5 and of four independent experiments for DNA5 in the presence of its preferred metal cation Mn^{2+}. *Error bars* represent the standard deviation. Linear regression analysis showed a correlation coefficient of 0.997 for RNA5 and 0.986 for DNA5.

mine the reason(s) for the different nuclease efficiencies of Sfn-His, we determined the Michaelis-Menten constant (K_m), maximal velocity (V_{max}), and the apparent rate of catalysis (k_{cat}) for comparable RNA and DNA substrates (Fig. 7). The K_m of RNA5 (K_m = 1.56 μM) is essentially identical to that of DNA5 (K_m = 1.51 μM). In contrast, the V_{max} and k_{cat} values were ~4-fold higher for RNA5 (V_{max} = 0.015 μM min^{-1}, k_{cat} = 1.5 min^{-1}) than for DNA5 (V_{max} = 0.004 μM min^{-1}, k_{cat} = 0.39 min^{-1}). Our K_m value for RNA5 is consistent with the micromolar range reported for *E. coli* Orn on $p(A)_5$ single stranded RNA substrates (19).

DISCUSSION

The human *SFN* gene belongs to the *yjeR/ORN* subfamily of the 3′ to 5′ exonuclease superfamily previously described (1, 5). Some of the members of this superfamily, most notably the Werner syndrome gene product and the polymyositis-scleroderma overlap syndrome 100-kDa autoantigen (PM-Scl 100), are associated with human disease (21, 22). While a connection of *SFN* to a specific human disease is not obvious, the ubiquitous expression of its transcript may suggest a general and essential role for the encoded protein in mammalian cells. Notably, the human *SFN* gene maps to chromosome position 11q23.1–11q23.2, a region that undergoes translocation events in several leukemias (23, 24), although none of these translocation breakpoints have been finely mapped to *SFN*.

The unique NH_2 terminus of hSfnα contains a consensus cleavage site pattern for mitochondrial processing proteases (17), whereas both Sfnα and Sfn possess a nuclear targeting signal in their COOH-terminal domains. This observation suggests that an alternative splicing event has evolved to give rise to a mitochondrial (Sfnα) and a nuclear (Sfn) version of the Ynt20/Orn equivalents in human. However, at present, the genomic DNA sequence of the entire *hSFN* gene is not available, and thus the alternatively spliced products cannot be confirmed. Notably, in the mouse EST data base, there are also multiple transcripts, suggesting that alternatively spliced murine *SFN* transcripts exist as well. Future studies will need to address whether there is tissue-specific expression of the mRNA splice variants found in mammals. Interestingly, while *S. cerevisiae* maintains a homologue to *hSFN*α (called *YNT20* or *Rex2*), we were unable to find a homologue to *SFN* in the NCBI data base using the Blastp search program.

A knockout of the *E. coli SFN* homologue, the *ORN* gene, results in cellular lethality (7). An *E. coli* temperature-sensitive mutant is not only lethal at the nonpermissive temperature (*i.e.* where Orn is inactive), but accumulates small oligoribonucleotides, indicating that Orn maintains an essential activity to degrade RNA (8). Three possibilities were provided as to why *ORN* deletion mutants are inviable (8): 1) accumulation of oligoribonucleotides results in a depletion of cellular mononucleotides; 2) accumulated oligoribonucleotides inhibit certain enzymes and interfere with essential metabolic processes; or 3) Orn has an additional unknown function that is responsible for the growth cessation. Since Sfn is capable of degrading both RNA and DNA, we propose that the human protein operates to remove not only short RNAs (likely resulting from RNA degradation processes) (25, 26), but also short single stranded DNAs that might arise as products of DNA repair and recombination. Thus, by extension (8), Sfn would function globally to recycle nucleotides or to remove nucleic acids that may interfere with essential cellular processes. How the nuclease activities of the Orn/Ynt20 proteins would function in mitochondrial DNA escape, a process that involves

translocation of mitochondrial DNA to the nucleus (9) and is possibly linked to cellular aging or senescence (27), is presently unclear. Furthermore, although unlikely, whether the NH_2-terminal differences between these proteins affect substrate specificity will need to be determined.

Not surprisingly, Sfn-His, which shares ~50% identity in amino acid sequence to its bacterial counterpart, has very similar biochemical properties to Orn (5). In particular, both proteins degrade short RNAs in a 3′ to 5′ direction, prefer Mn^{2+} as its metal cofactor, and have an enzymatic activity temperature optimum of ~50 °C. However, gel filtration chromatography suggests that the human Sfn-His fusion protein is a tetramer, in contrast to *E. coli* Orn, which exists as a homodimer (6), suggesting that these proteins may maintain differing activities as well.

The nuclease activities of both Sfn and Orn (19) are inversely proportional to the length of the single stranded substrate. Furthermore, previous kinetic data of *E. coli* Orn has shown that the K_m values for short and long single stranded RNAs are similar (19). We observed that Sfn has a similar K_m for short single stranded RNAs and DNAs, but that such short RNAs are degraded ~4-fold more efficiently than DNAs. Such experimental observations raise questions of how the substrate length or the nucleic acid chemistry influences Sfn/Orn enzymatic activity. To answer these questions will require determining which step(s) is influenced by nucleotide length or nucleic acid composition. High resolution structural data of Sfn, alone and in complex with RNA and DNA, would also shed light on the recognition and catalytic mechanisms of these proteins. Lastly, our studies emphasize that future experiments should pay particular attention to the potential range of substrate diversity recognized by the other 3′ to 5′ exonuclease superfamily members.

REFERENCES

1. Moser, M. J., Holley, W. R., Chatterjee, A., and Mian, I. S. (1997) *Nucleic Acids Res.* **25,** 5110–5118
2. Bernad, A., Blanco, L., Lazaro, J. M., Martin, G., and Salas, M. (1989) *Cell* **59,** 219–228
3. Koonin, E. V., and Deutscher, M. P. (1993) *Nucleic Acids Res.* **21,** 2521–2522
4. Joyce, C., and Steitz, T. (1994) *Annu. Rev. Biochem.* **63,** 777–822
5. Koonin, E. V. (1997) *Curr. Biol.* **7,** R604–R606
6. Zhang, X., Zhu, L., and Deutscher, M. P. (1998) *J. Bacteriol.* **180,** 2779–2781 **AQ: D**
7. Deutscher, M. P. (1993) *J. Bacteriol.* **15,** 4577–4578
8. Ghosh, S., and Deutscher, M. P. (1999) *Proc. Natl. Acad. Sci. U. S. A.* **96,** 4372–4377
9. Shafer, K. S., Hanekamp, T., White, K. H., and Thorsness, P. E. (1999) *Curr. Genet.* **34,** 438–448
10. Niyogi, S. K., and Datta, A. K. (1975) *J. Biol. Chem.* **250,** 7307–7312
11. Wilson, D. M., III, Carney, J. P., Coleman, M. A., Adamson, A. W., Christensen, M., and Lamerdin, J. E. (1998) *Nucleic Acids Res.* **26,** 3762–3768
12. Burgess, R. R. (1991) *Methods Enzymol.* **208,** 3–10
13. Pace, C. N., Vajdos, F., Fee, L., Grimsley, G., and Gray, T. (1995) *Protein Sci.* **4,** 2411–2423
14. Erzberger, J. P., and Wilson, D. M., III (1999) *J. Mol. Biol.* **290,** 447–457
15. Derbyshire, V., Grindley, N. D., and Joyce, C. M. (1991) *EMBO J.* **10,** 17–24
16. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual,* 2nd Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
17. Gavel, Y., and von Heijne, G. (1990) *Protein Eng.* **4,** 33–40
18. Hicks, G. R., and Raikhel, N. V. (1995) *Annu. Rev. Cell Dev. Biol.* **11,** 155–188
19. Datta, A. K., and Niyogi, S. K. (1975) *J. Biol. Chem.* **250,** 7313–7319
20. Lee, B. I., and Wilson, D. M., III (1999) *J. Biol. Chem.* **274,** 37763–37769
21. Shen, J. C., Gray, M. D., Oshima, J., Kamath-Loeb, A. S., Fry, M., and Loeb, L. A. (1998) *J. Biol. Chem.* **273,** 34139–34144
22. Bluthner, M., and Bautz, F. A. (1992) *J. Exp. Med.* **176,** 973–980
23. Ziemin, L., van der Poel, S., McCabe, N. R., Gill, H. J., Espinosa, R., III, Patel, Y., Harden, A., Rubinelli, P., Smith, S. D., Le Beau, M. M., Rowley, J. D., and Diaz, M. O. *Proc. Natl. Acad. Sci. U. S. A.* **88,** 10735–10739
24. Chen, Z., Brand, N. J., Chen, A., Chen, S. J., Tong, J. H., Wang, Z.-Y., Waxman, S., and Zelent, A. (1993) *EMBO J.* **12,** 1161–1167
25. Morrissey, J. P., and Tollervey, D. (1995) *Trends Biochem. Sci.* **20,** 78–82
26. Jacobson, A., and Peltz, S. W. (1996) *Annu. Rev. Biochem.* **65,** 693–739
27. Shafer, K. S., Hanckamp, T., White, K. H., and Thorsness, P. E. (1999) *Curr. Genet.* **36,** 183–194 **AQ: G**

THE HUMAN INTEFERON AND ESTROGEN-REGULATED *HEM45/ISG20* GENE PRODUCT DEGRADES SINGLE STRANDED RNA AND DNA *IN VITRO*

by

Lam H. Nguyen and David M. Wilson III*

Molecular and Structural Biology Division

Lawrence Livermore National Laboratory

P.O. Box 808, L-441

Livermore, CA 94551

*Corresponding author: wilson61@llnl.gov

925-423-0695

925-422-2282 (Fax)

Running Title: Hem 45/Isg20 is an exonuclease

SUMMARY

The human *HEM45/ISG20* gene is directly regulated by estrogen hormone and interferon. Amino acid sequence comparison suggests that this gene encodes a protein with homology to the 3' to 5' exonuclease superfamily that includes RNases (such as RNase T and D), the proofreading domains of DNA polymerase I family. We provided here direct biochemical evidence that *HEM45/ISG20* gene product acts as a 3' to 5' exonuclease *in vitro*. This protein degrades RNA and single stranded DNA with ~35 fold less efficiency. Hem45/Isg20 prefers Mn^{2+} as metal cofactor and displays pH optima of ~7.0. Mutation of a conserved aspartate (D94) to glycine reduces both nuclease activities of Hem45/Isg20. In addition to RNase L, Hem45/Isg20 is the second known interferon-regulated RNase. Hem45/Isg20 is located in PML nuclear bodies which are known sites of transcription (and replication) of not only oncogenic DNA viruses but also cellular hormonal-dependent RNA polymerase II transcription. We proposed that the biological role of Hem45/Isg20 is to degrade not only viral RNAs as part of the cellular interferon anti-viral response, but also estrogen regulated cellular mRNAs as part of the down regulation of estrogen signaling process.

INTRODUCTION

HEM45 (HeLa Estrogen Modulated, band 45) gene, also called *ISG20* (Interferon Stimulated Gene product of 20 kDa), is ubiquitously expressed as a ~0.7 kb transcript at low level in numerous immortal human cell lines, in various human tissues, and also in rat immature uterus (1, 2). Transcription of the *HEM45/ISG20* gene is directly regulated by estrogen hormone (2). Estrogen hormones cause *HEM45/ISG20* mRNA level to increase ~20-fold in human HeLa cells expressing the estrogen receptor alpha gene and to increase maximally 8-fold in rat uterus six hours after estrogen treatment (2). Estrogens are important in normal uterine development (2).

The *HEM45/ISG20* gene is also induced directly by interferons at the transcription level (1). Interferons are a family of secreted cellular proteins that are involved in the regulation of cell proliferation, cellular differentiation, immune response, and antiviral response (reviewed in 3). The *HEM45/ISG20* gene has been shown to localize to chromosome position 15q26 by in situ hybridization (4), while the Hem45/Isg20 proteins, like some other interferon-induced proteins, are located in spherical nuclear bodies

termed promyoletic (PML) oncogenic domains (PODs) [1]. PODs (also known as nuclear domain 10 or Kr bodies) are distinct from other subnuclear domains such as nucleolus and splice sites (5). PODs are composed of large multiprotein complexes associated with the nuclear matrix (5). They represent preferential targets for viral infection (6) and interferon action, and their substructure integrity are compromised in certain human diseases including leukemia and viral infections (6).

HEM45/ISG20 shows amino acid sequence homology to the carboxyl terminus of *XPMC2* gene of human and of *Xenopus laevis* (1, 7). *XPMC2* appears to be a negative regulator of cell division involving in the control mechanism that prevents the onset of mitosis until DNA replication is completed. *XPMC2* is able to prevent mitotic catastrophe in *Schizosaccharomyces pombe* strain that lack inhibitors of the G2 to M transition of the cell cycle (8). Xpmc2 is localized to the nucleus of Xenopus eggs and was found to be a substrate for MPF (8).

The *HEM45/ISG20* gene was recently placed into the 3′ to 5′ exonuclease superfamily (9) because it contains conserved exonuclease motifs. Homology within the superfamily is centered

around three conserved exonuclease motifs termed Exo I, II and III (10). The crystal structure of the Klenow fragment of *Escherichia coli* DNA polymerase I and complementing site-directed mutagenesis studies indicate that the Exo I, II, and III motifs are clustered around the active site and contain four conserved negatively charged residues that are critical for coordinating the two metal ions involved in phosphodiester bond cleavage (reviewed in 11). This superfamily includes RNases (such as RNaseT and D), the proofreading domains of pol I-family DNA polymerases, and DNases that exist as independent proteins (such as *Escherichia coli* Exo1) or as domains within larger polypeptides (such as a region within the Werner's syndrome helicase) [9, 10, 11].

We reported here the biochemical evidence that Hem45/Isg20 is a 3' to 5' exonuclease with a strong preference for RNAs over single stranded DNAs. We discussed the biological role Hem45/Isg20 may play in the cellular interferon-dependent antiviral response and estrogen hormonal signaling process.

EXPERIMENTAL PROCEDURES

Buffers and Reagents-All reagents were purchased from Sigma unless otherwise indicated. Restriction enzymes were purchased from New England Biolabs. Labeled nucleotides were from Amersham. Spectrophotometric-grade glycerol was obtained from Fisher. pETs vectors were from Novagen (Madison, WI). DNA oligos were obtained from Operon Biotechnologies (Alameda, CA). Synthetic RNA was obtained from Dharmacon Research (Boulder, CO). L buffer (Lysis buffer) contains 50 mM Tris-HCl pH 7.0, 50 mM NaCl, 20% glycerol, 20 mM β-mercaptoethanol, and 1 mM phenylmethylsulfonylflouride (PMSF). W5 buffer is 50 mM Tris-HCl pH 7.9, 500 mM NaCl, 20% glycerol, 5 mM imidazole, 10 mM β-mercaptoethanol, and 1 mM PMSF. W20 is 50 mM Tris-HCl pH 7.9, 500 mM NaCl, 20% glycerol, 20 mM imidazole, 10 mM β-mercaptoethanol, and 1 mM PMSF. E250 buffer is 50 mM Tris-HCl pH 7.9, 250 mM NaCl, 20% glycerol, 250 mM imidazole, 10 mM β-mercaptoethanol, and 1 mM PMSF. All pH values were determined at 21°C.

Purification of recombinant Hem45/Isg20-His protein-The *HEM45/ISG20* coding region was PCR amplified using primers

NCO5´HEM45 (5´-GCATGCAGAATTCGGCACGAGCTCTGA-3') and HIND3´HEM45 (5´-GCAGTAAGCTTTCAGTCTGACACAGCCAG-3'), partially digested with NcoI and HindIII, and sub-cloned after digestion into the NcoI and HindIII sites of pET28d (Novagen, Madison, WI) to generate pHem45/Isg20-His. This construct allows for expression of six histidine residues on the carboxyl terminus of the Hem45/Isg20 protein under the control of a T7 RNA polymerase promoter. pHem45/Isg20-His plasmid was sequenced (12) and no PCR errors were found.

pHem45/Isg20D94G plasmid contains the *HEM45/ISG20* gene with D94G mutation in the ExoII exonuclease motif (9). A 145 bp PstI fragment containing this mutation was gel isolated from a PstI digest of pGST-ISG20ExoII (obtained from Dr. Mechti) and subcloned into the internal PstI sites within the *HEM45/ISG20* protein coding region of pHem45/Isg20-His plasmid.

The pHem45/Isg20-His or pHem45/Isg20D94G plasmid was transformed into BL21(DE3)/pLysS *E.coli* strain (Novagen). An overnight culture of 100 ml was grown at 37°C in LB (1% bacto-tryptone, 0.5% bacto-yeast extract, and 1% NaCl) with 50 μg/ml kanamycin and 25 μg/ml chloramphenicol. The overnight culture

was used to inoculate 2 liter of the same medium. This culture was grown at 37°C with vigorous aeration until the A_{550nm} was ~0.6. Isopropyl-1-thio-β-galactopyranoside (final concentration of 1 mM) was then added to induce Hem45/Isg20-His protein expression. After 4 hours of induction, cells were harvested, and cell pellets were frozen at –20°C. pHem45/Isg20-His plasmid, but not pHem45D94G plasmid, causes ~2-3 fold decrease of growth rate of BL21(DE3) cells even in the presence of plysS plasmid at the uninduced state (data not shown), suggesting a role for nuclease activity in growth efficiency. The cell pellet was resuspended in 40 ml of L buffer. The cell suspension was sonicated using a Misonix XL sonicator by two 1 minute bursts with a microtip at the maximum setting. The cell lysate was centrifuged at 27,000g for 30 min at 4°C. To the supernatant, polyethyleneimine (PEI) was added slowly with constant stirring to a final concentration of 0.15% to remove nucleic acids (13). The suspension was centrifuged and to the supernatant, NaCl was added to a final concentration of 500 m M and imidazole was added to a final concentration of 5 mM. The eluant was then incubated with 2 ml of Ni-NTA resin (Qiagen, Santa Clarita, CA) for 1 hr at 4°C with gentle rocking. This suspension was

centrifuged, and the resin was washed 2 times each with 20 ml W5 buffer, followed by 2 washes each with 20 ml of W20 buffer. Hem45/Isg20-His proteins were then eluted from the Ni-NTA resin 4 times, each with 2 ml of E250 buffer, dialyzed overnight against L buffer containing 0.1 mM DTT and 50% glycerol, and stored at −20°C. Concentrations of protein solutions were determined by measuring the absorbance at 280nm and using the theoretical molar extinction coefficient E_{280nm} of 13,325 $M^{-1}cm^{-1}$ calculated for Hem45/Isg20-His (14). From 2 liter of induced bacterial culture, we obtained ~1 mg of Hem45/Isg20-His fusion protein (Figure 1).

Nuclease assays- Various DNA and RNA5 substrates were labeled at the 5' end by T4 polynucleotide kinase with γ-P^{32}-ATP (15). RNA5 is 5'-GAUCG-3'. DNA5 is 5'-GATCG-3'. Uniformly labeled RNA substrates used in Figure 3 were synthesized with T7 RNA Polymerase and α-P^{32}-GTP according to the instructions of the manufacturer (Gibco-BRL). Supercoiled pET-28d was used as the promoter DNA template to produce 343 base long RNA with a 15 base stem and 6 base loop structure (TΦ transcription terminator) at the 3' end of the RNA (16). For the 210 base long RNA, pET-28d plasmid that has been linearized with XhoI was used as the T7

promoter template, with the XhoI restriction site is located between the T7 promoter and the TΦ transcription terminator.

Labeled nucleic acid substrate (200-400 fmol) was incubated with 40-1000 fmol of Hem45/Isg20-His protein at 37°C in 10 µl of 50 mM HEPES-KOH pH 7.0, 10% glycerol, 50 mM NaCl, 1 mM $MnCl_2$, 0.01% Triton X-100, and 1 mM DTT. The reactions were stopped with 10 µl of 80% formamide dye solution, heated at 80°C for 3 min, and then fractionated on a 20% or a 22.5% 8 M urea polyacrylamide gel. Visualization of the labeled substrate on the gels was also achieved using a Molecular Dynamics (Sunnyvale, CA) STORM 860 Phosphorimager and quantitative analysis was performed using Molecular Dynamics ImageQuant v1.11 software. For the pH titrations, different pHs were maintained with a buffer mixture of constant ionic strength composed of 25 mM of acetic acids and MES, and 50 mM of Trizma base, instead of 25 mM HEPES-KOH pH 7.0 (17). The specific activity is defined as 1U/µg of Hem45/Isg20-His. 1U is 1pmol of RNA5 or DNA5 substrate being degraded to mononucleotide form in 1 minute at 37°C.

RESULTS

Hem45/Isg20-His is a 3' to 5' exonuclease-The *HEM45/ISG20* cDNA codes for a 181 amino acid protein of 20,363 daltons with a theoretical pI of 9.5. Hem45/Isg20 belongs to the 3' to 5' exonuclease superfamily (9) based on conservation of the three characteristic sequence motifs ExoI, II, and III (10, 11).

To examine experimentally for nuclease activity, *HEM45/ISG20* gene was subcloned into the pET28d plasmid to produce a C-terminal-tagged Hem45/Isg20-His fusion protein, and the recombinant protein was purified as described in Experimental Procedures (Figure 1).

As shown in Figure 2, Hem45/Isg20-His is capable of degrading RNA5, DNA5, and the single stranded DNA flap of the 3' flap DNA substrate, stopping at the double stranded DNA junction. Also, Hem45/Isg20-His is ~35-fold more active on single stranded RNA than DNA (Figure 2), with a specific activity for RNA5 of 4.21 ± 0.80 U/μg, and for DNA5 substrate of 0.12 ± 0.01 U/μg. We observed degraded 4 mer products prior to detecting mononucleotide products in a time course with single stranded

RNA5 or DNA5 substrate (Figure 2), consistent with Hem45/Isg20-His as a 3' to 5' exonuclease.

We also examined Hem45/Isg20-His RNase activity on longer RNA substrate with or without a stem-loop structure at the 3' end. As shown in Figure 3, Hem45/Isg20-His is capable of degrading not only short RNA such as RNA5, but also 210 base long RNA. However, the presence of a stem-loop structure at the 3' end of the RNA substrate cause a ~40 fold reduction in the RNase activity of Hem45/Isg20-His at the 30 or 60 time point (Figure 3). Such data is consistent with the above time course data indicating that Hem45/Isg20-His is a 3' to 5' exonuclease. The data also suggest that Hem45/Isg20-His exonuclease activity on RNA is processive and not distributive since, in a time course, we did not detect many of intermediate size products between the starting RNA substrate and the final mononucleotide products (Figure 3).

We subsequently examined Hem45/Isg20 nuclease activity on other DNA structures known to be products/substrates of DNA repair processes. The various double stranded DNA substrates we used are described in Ref. 18. We did not detect any protein-dependent degradation of 42 bp double-stranded DNAs of gapped,

nicked, or flushed structures (data not shown). Neither did we detect any degradation of 5' flap double stranded DNA structure (data not shown). However, we did detect degradation of the single stranded DNA region of a 3' overhang DNA structure (data not shown). Such data is consistent with the results of Figure 2, indicating that Hem45/Isg20-His degrades single stranded RNA and DNA, but not double stranded DNA structures.

To confirm that the observed DNase and RNase activities are intrinsic to the Hem45/Isg20-His protein, a catalytically inactive Hem45/Isg20-His mutant was constructed. Amino acid residue Aspartate^{94} was selected because an aspartate to alanine mutation at the structurally equivalent residue in *Escherichia coli* DNA polI causes a large decrease in 3' to 5' exonuclease activity (18). The Hem45/Isg20D94G mutant was purified using the same procedure as described for wild type Hem45/Isg20-His protein (Figure 1). When compared to wild type Hem45/Isg20-His, the Hem45/Isg20D94G mutant displayed a ~90% reduction in degrading both RNA5 and DNA5 substrates (Figure 4). These data are

consistent with both nuclease activities being intrinsic to the Hem45/Isg20-His protein, and not being the result of contaminants. *Optimization of the exonuclease activity of Hem45/Isg20-His fusion protein*-Using a 5' end labeled RNA5 as the substrate, we determined the divalent metal cation preference of human Hem45/Isg20-His protein. At metal concentrations of 0.05, 0.5, and 5 mM, only Mn^{2+} had any stimulatory effect on Hem45/Isg20-His exoribonuclease activity, while in the presence Mg^{2+}, Ca^{2+}, Cu^{2+}, Zn^{2+}, and Ni^{2+}, there was no detectable degradation of the RNA5 substrate (Table 1). The DNase activity, as determined with the 3' flap DNA substrate, also showed the same metal preference (data not shown). In addition, both the RNase and DNase activities of Hem45/Isg20-His show the same pH profile, with the optimal pH of 7.0 (Figure 5 and data not shown).

DISCUSSION

The human *HEM45/ISG20* gene is categorized as a member of the 3' to 5' exonuclease superfamily (9, 10, 11) based on amino acid sequence comparisons. We presented here direct biochemical evidence that Hem45/Isg20 is a 3' to 5' exonuclease with strong preference for single stranded RNA over single stranded DNA. Our studies here, along with the studies on RNase T and Orn proteins (20, 21), emphasize that future experiments should pay particular attention to the potential range of substrate diversity recognized by the other 3' to 5' exonuclease superfamily members.

Some of the members of this superfamily, most notably the Werner syndrome gene product (WRN) and the polymyositis-scleroderma overlap syndrome 100 kDa autoantigen (PM-Scl 100), are associated with human disease (9). While a connection of Hem45/Isg20 to a specific human disease is not known, RNases are now known to control biological processes ranging from splicing to organogenesis (reviewed in 22). Growth regulatory molecules such as angiogenin, neurotoxins, and plant allergens have RNase activity or significant homology to known RNases (reviewed in 22).

While the biological significance of the weak DNase activity of Hem45/Isg20 is unclear, its RNase activity may play a role in the cellular interferon antiviral response. *HEM45/ISG20* is expressed at low level in all cell types studied so far (1, 2) and is up regulated at the transcription level by interferons (1). Transient transfection experiments indicated that Hem45/Isg20 is located in the PML nuclear bodies (PODs) near the nuclear matrix (1). The PODs are distinct from other known nuclear substructures such as nucleolus and splice sites. Such PODs are known sites of transcription and replication of oncogenic DNA viruses such as Simian Virus SV40, Adenovirus, and Herpes Simplex Virus (6). These experimental observations, taken together, suggest that Hem45/Isg20 is part of the cellular interferon defense mechanism against viral infection. We proposed that Hem45/Isg20, given its biochemical activity of degrading RNAs, is involved in removing viral mRNAs to prevent viral gene expression and thereby viral propagation.

Hem45/Isg20 is the second known RNase regulated by interferon. RNase L, the first interferon-dependent RNase, is also constitutively expressed at a low level in various cell lines like Hem45/Isg20 (reviewed in 3). However, in contrast to

Hem45/Isg20, RNase L is active only in the presence of its ligand effector, 2'5'linked poly-A (reviewed in 3). It is conceivable that there may be other protein or ligand inhibitors that can complex with Hem45/Isg20 in order to regulate its RNase activity since it is thought that RNases are typically present in minute quantities bound with inhibitors or lying latent in cells awaiting an activator (reviewed in 22). Alternatively, Hem45/Isg20 activity can be regulated by post-translational modification. However, we suggest that Hem45/Isg20 RNase activity is probably regulated by localizing it to a particular region of the cell such as PODs since, in transient transfection experiments, overexpressed Hem45/Isg20 are localized to the PODs and such overexpression did not appear to affect cell growth (1).

The estrogen hormone, like interferon, is also capable of transcriptionally up regulated the *HEM45/ISG20* gene (2). Recent demonstrations of nascent RNA polymerase II transcripts and CBP (CREB binding proteins) transcription coactivators in PODs suggest a role for PODs in transcriptional regulation (23). Specifically, at least one POD associated protein, PML, has been shown to interact with CBP and dramatically stimulated hormonal nuclear receptor (such

as retinoid X receptor and glucocorticoid receptor) transcriptional activity (24). The hormonal receptors, when complexed with their respective hormone ligand, act as transcription factors. These results raise the possibility that PML may also stimulate estrogen nuclear receptor transcriptional activity. If this is the case, then it is conceivable that the function of Hem45/Isg20, induced by estrogen and localized to the PODs, is to down-regulate the estrogen-dependent transcriptional response by degrading mRNAs induced by estrogen. Consistent with this hypothesis, the increase of *HEM45/ISG20* mRNA level in immature uterus in the presence of estrogen occurred later than those of estrogen early response genes such as c-jun, c-fos, and zif-268 (2). c-jun, c-fos, and zif-268 are transcription factors that are thought to mediate the early stage of the uterine response to estrogens (2). Future experiments will be needed to test these hypotheses about the precise role of Hem45/Isg20 in the cellular interferon-dependent antiviral response and estrogen hormonal signaling pathway.

ACKNOWLEDGEMENTS

This work was carried out under the auspices of the U. S. Department of Energy by Lawrence Livermore National Laboratory under contract number W-7405-ENG-48 and supported by NIH (CA79056) and U. S. Army USAMRMC (BC980514) grants to DMWIII.

REFERENCES

1. Gongora, C., David, G., Pintard, L., Tissot, C., Hua, T. D., Dejean, A., and Mechti, N. (1997) *J. Biol. Chem.* **272**, 19,457-19,463.
2. Pentecost, B. T. (1998) *Steroid Biochem. Mol. Biol.* **64**, 25-33.
3. Lengyel, P. (1993) *Proc. Natl. Acad. Sci. U. S. A.* **90**, 5,893-5,895.
4. Mattei, M. G., Tissot, C., Gongora, C., and Mechti, N. (1997) *Cytogenet. Cel. Genet.* **79**, 286-287.
5. Seeler, J. S. and Dejean, A. (1999) *Curr. Opin. Genet. Dev.* **9**, 362-367.
6. Maul, G. G. (1998) *BioEssays* **20**, 660-667.
7. Kwiatkowska, J., Slomski, R., Jozwiak, S., Short, M. P., and Kwiatkowski, D. J. (1997) *Genomics* **44**, 350-354.
8. Su, J. Y and Maller, J. L. (1995) *Mol. Gen. Genet.* **246**, 387-396.
9. Moser, M. J., Holley, W. R., Chatterjee, A., and Mian, I. S. (1997) *Nucleic Acids Res.* **25**, 5,110-5,118.
10. Moser, M. J., Holley, W. R., Chatterjee, Bernad, A., Blanco, L., Lazaro, J. M., Martin, G., and Salas, M. (1989) *Cell* **59**, 219-228.
11. Joyce, C. and Steitz, T. (1994) *Annu. Rev. Biochem.* **63**, 777-822.

12. Wilson III, D. M., Carney, J. P., Coleman, M. A., Adamson, A. W., Christensen, M., and Lamerdin, J. E. (1998) *Nucleic Acids Res.* **26**, 3,762-3,768.

13. Burgess, R. R. (1991) *Methods Enzymol.* **208**, 3-10.

14. Pace, C. N, Vajdos, F., Fee, L., Grimsley, G., and Gray, T. (1995) *Protein Sci.* **4**, 2,411-2,423.

15. Erzberger, J. P. and Wilson III, D. M. (1999) *J. Mol. Biol.* **290**, 447-457.

16. Studier, F. W., Rosenberg, A. H., Dunn, J. J., and Dubendorff, J. W. (1990) *Methods. Enzymol.* **185**, 60-89.

17. Ellis, K. J. and Morrison, J. F. (1982) *Meth. Enzymol.* **87**, 405-426.

18. Lee, B. I. and Wilson III, D. M. (1999) *J. Biol. Chem.* **274**, 37,763-37,769.

19. Derbyshire, V., Grindley, N. D., and Joyce, C. M. (1991) *EMBO J.* **10**, 17-24.

20. Viswanathan, M., Dower, K. W., and Lovett, S. T. (1998) J. *Biol. Chem.* **273**, 35,126-35,131.

21. Nguyen, L. H., Erzberger, J. P., Root, J., and Wilson, D. M. III (2000) *J. Biol. Chem.* Manuscript accepted.

22. Schein, C. H. (1997) *Nat. Biotechnol.* **15**, 529-535.

23. Doucas, V., Tini, M., Egan, D. A., and Evans, R. M. (1999) *Proc. Natl. Acad. Sci. U. S. A.* **96**, 2627-2632.

24. LaMorte, V. J., Dyck, J. A., Ochs, R. L., and Evans, R. M. (1998) *Proc. Natl. Acad. Sci. U. S. A.* **95**, 4991-4996.

25. Diaz, G. A., Rong, M., McAllister, W. T., and Durbin, R. K. (1996) *Biochemistry* **35**, 10,837-10,843.

FIGURE LEGENDS

Figure 1. Purity of Hem45/Isg20-His and Hem45/Isg20ExoII proteins. A 15% polyacrylamide SDS gel containing fractions of the various purification steps. Lanes 1-5 are Hem45/Isg20-His; lanes 6-10, Hem45/Isg20D94G; and lane 11, protein markers. Lanes 1 and 10, whole cell extract; lanes 2 and 9, whole cell supernatant; lanes 3 and 8, 0.15% PEI supernatant; lanes 4 and 7, flow-through of the Ni-affinity chromatography; and lanes 5 and 6, Ni-affinity purified fraction. Protein marker sizes in kDa are indicated on the right. The gel was stained with Coomassie Blue R250.

Figure 2. Time course of Hem45/Isg20-His exonuclease activity. Hem45/Isg20-His degrades single stranded RNAs and DNAs. All substrates were labeled at the 5' end. 40 nM of RNA5 or DNA substrates (DNA5 or 3' flap) was incubated with 20 nM or 200 nM of Hem45/Isg20-His protein, respectively, for various times at 37°C. RNA5 or DNA5 was labeled at the 5' end. The top DNA strand containing the flap was labeled at the 5' end also. Lanes 1-6 are reactions with RNA5; lanes 7-12, DNA5; lanes 13-18, 3' flap DNA substrate of 42 mer with 10 nucleotide long flap at the 3' end. Lanes 1, 7, and 13 are the no protein controls. Lanes 2, 8, and 14, 2

min incubations with Hem45/Isg20-His; lanes 3, 9, and 15, 4 min incubations; lanes 4, 10 and 16, 8 min incubations; lanes 5, 11, 17, 16 min incubations; and lanes 6, 12, and 18, 32 min incubations.

Figure 3. Hem45/Isg20-His nuclease activity on long RNA substrates with or without a stem-loop structure at the 3' end. Lanes 1-4 are reactions with uniformly labeled 343 base long RNA with the TΦ transcription terminator stem-loop structure at the 3' end; lanes 5-8, uniformly labeled 210 base long RNA; and lanes 9-12, RNA5 labeled at the 5' end. Lanes 1, 5, and 9 are the no protein controls. All other lanes are reactions with 10 nM of Hem45/Isg20-His. Lanes 2, 6, and 10 are 30 min reactions at 37°C; lanes 3, 7, and 11, 60 min reactions; and lanes 4, 8 and 12, 90 min reactions. On the right of the figure, band A is unincorporated α-P^{32}-GTP; band B, Hem45/Isg20-His degraded mononucleotide products; band C, RNA5 substrate; band D, 210 base long RNA; and band E, 343 base RNA with a 3' end stem-loop structure. Signals that are located immediately on top of band B in lanes 5-8 are probably abortive initiation RNA products from linearized DNA promoter template but not from supercoiled promoter template (25). Band E signal in

lanes 5-8 is due to incomplete linearization of the promoter template by XhoI.

Figure 4. The effect of a D94G mutation on the DNase and RNase activities of Hem45/Isg20-His. Exonuclease activity of Hem45/Isg20D94G mutant with RNA5 and 3' flap DNA substrate described in Figure 2. Lanes 1-5 are with RNA5, and lanes 6-10 are with 3' flap substrates. Lane 11 is DNA markers of 42 mer and 21 mer. The substrate concentration is 40 nM. Lane 1 is RNA5 undigested control; protein control; lane 2 and 3, 5 and 15 nM of wild type Hem45/Isg20-His (WT); lane 4 and 5, 5 and 15 nM of Hem45/Isg20D94G mutant (D94G mutant); lane 6, 3' flap DNA undigested control; lanes 7 and 8, 50 and 150 nM of WT enzyme; and lanes 9 and 10, 50 and 150 nM of D94G mutant enzyme. The reactions were incubated at 37°C for 20 min. Shown is a representative of three independent experiments.

Figure 5. pH optima of Hem45/Isg20-His exonuclease activity. 40 nM of 5'-labeled RNA5 was incubated with 2 nM of Hem45/Isg20-His for 10 minutes 37°C at various pHs. pH was maintained with a mixture of 25 mM acetic acid, 25 mM of MES, and 50 mM of TrisHCl (16). The x-axis is the pH. The y-axis is the

relative exonuclease activity. Maximal activity at pH 7 is arbitrarily designated as 1. Exonuclease activity is proportional to the amount of labeled RNA5 substrate that has been converted to the final mononucleotide product. Shown is a representative of three independent experiments.

Table 1. Metal cofactor requirement of Hem45/Isg20-His[a].

Metal	Concentration (mM)	Relative Activity
$MnCl_2$	5	0.85
	0.5	1
	0.05	0.15
$MgCl_2$	5,0.5, 0.05	<0.01
$CaCl_2$	5,0.5, 0.05	<0.01
$CuCl_2$	5,0.5, 0.05	<0.01
$ZnCl_2$	5,0.5, 0.05	<0.01
$NiCl_2$	5,0.5, 0.05	<0.01

[a]40 nM of labeled RNA5 substrate was incubated with 2 nM of Hem45/Isg20-His enzyme for 10 min at 37°C with different concentrations of various divalent metals shown. Mononucleotide products were then quantitated by Phosphorimager scan and used to represent RNase activity. The highest RNase activity was arbitrarily designated as 1. Shown is a representative of three independent experiments.

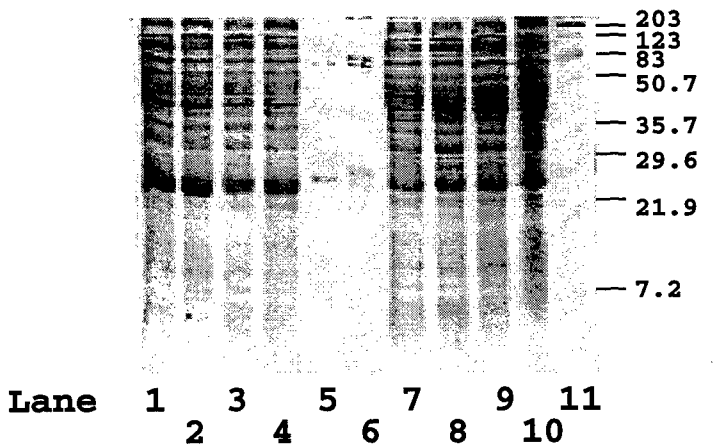


Figure 1

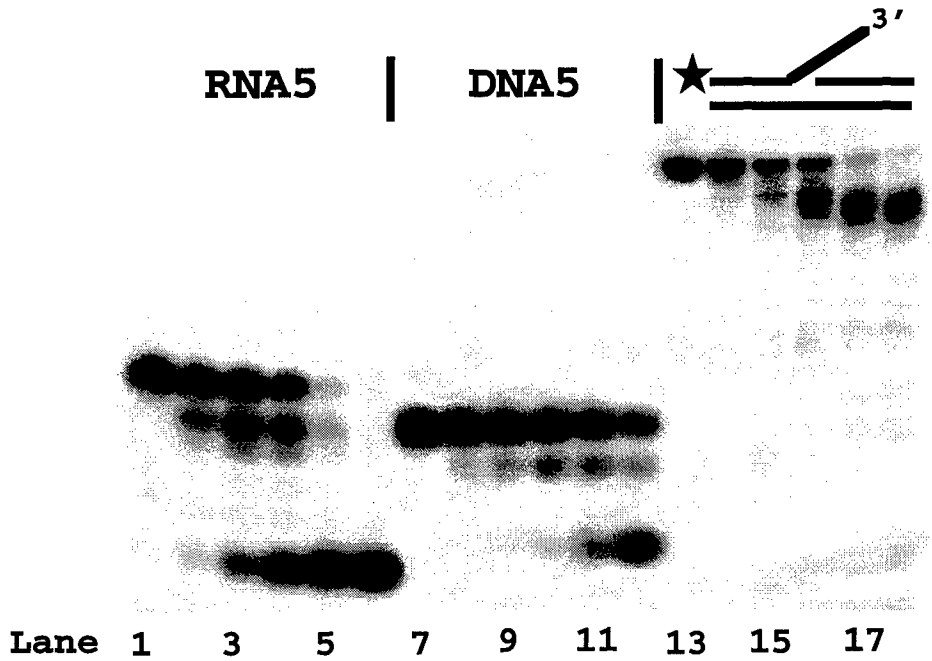


Figure 2

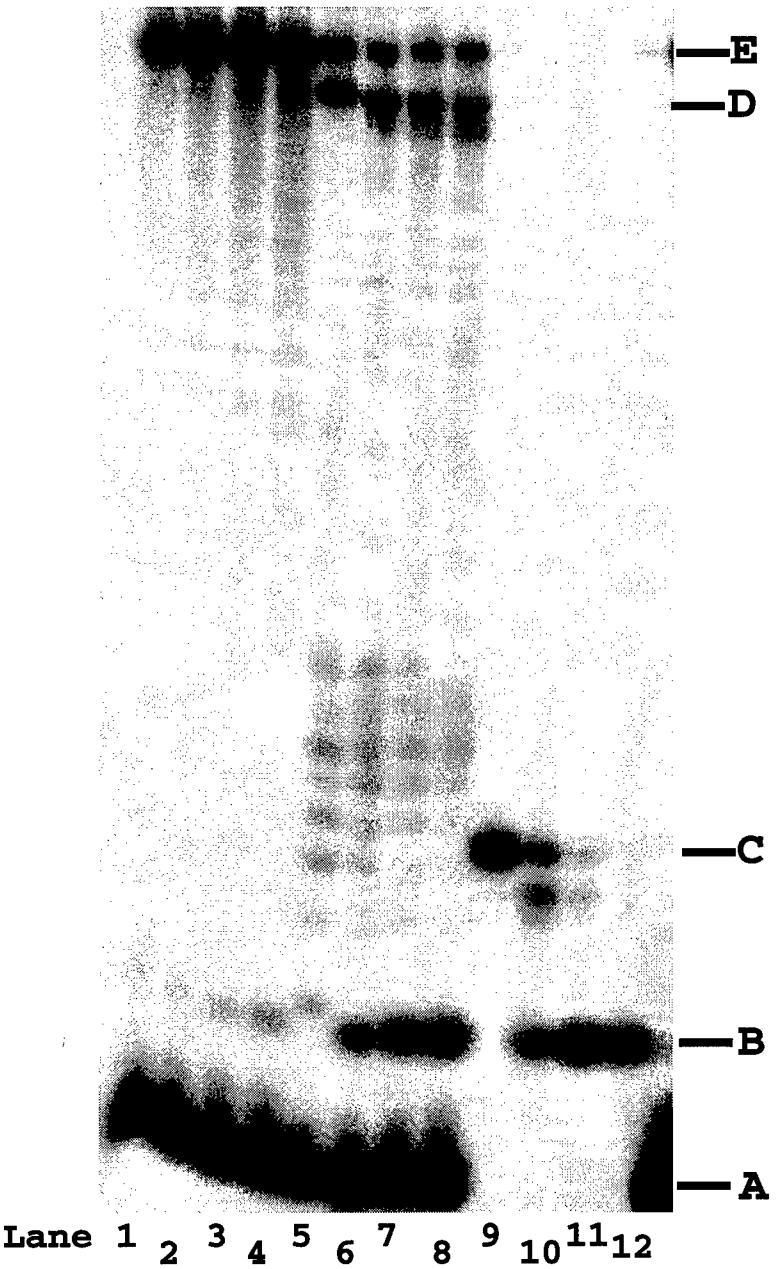


Figure 3

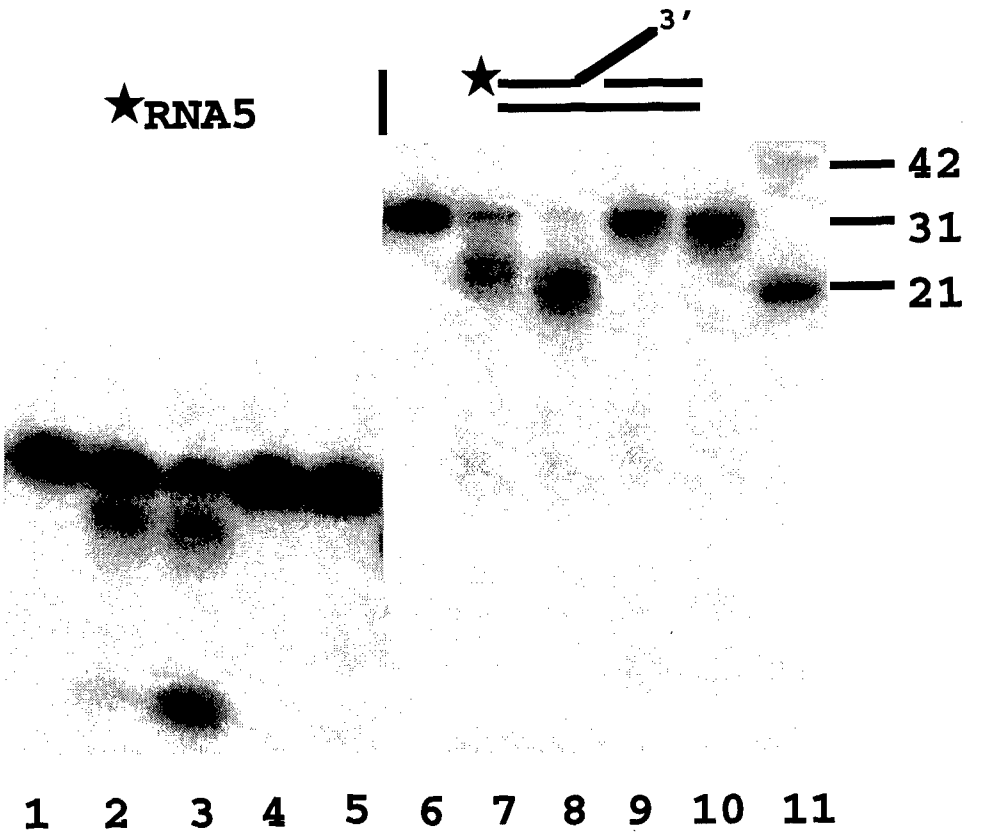


Figure 4

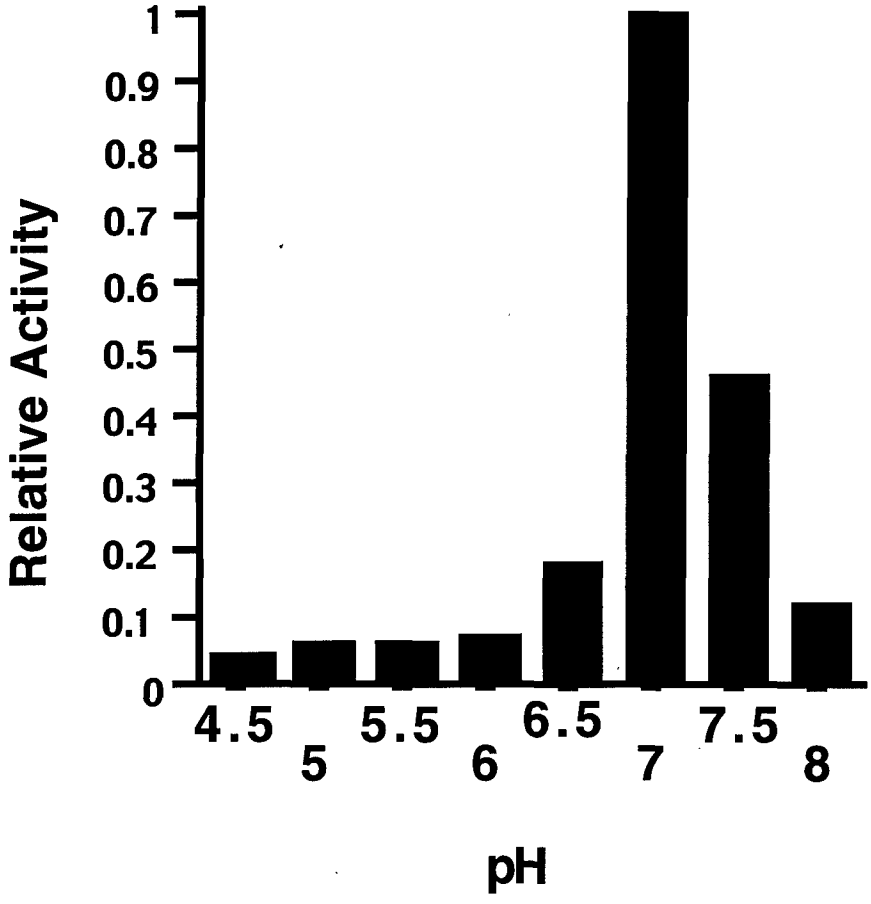

Relative Activity
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
4.5
5
5.5
6
6.5
7
7.5
8
pH

Figure 5